

2021-07

Temporal Variations in Landslide Distributions Following Extreme Events: Implications for Landslide Susceptibility Modeling

Jones, JN

<http://hdl.handle.net/10026.1/17466>

10.1029/2021jf006067

Journal of geophysical research. Earth surface

American Geophysical Union

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.



RESEARCH ARTICLE

10.1029/2021JF006067

Key Points:

- We use a 30-yr multi-temporal landslide inventory to quantify the temporal variations in landslide spatial distributions
- Landslide spatial distributions are found to vary significantly through time, particularly in years impacted by extreme events
- Temporal variation in landslide occurrence challenges the time-independent nature of typical landslide susceptibility models

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J. N. Jones,
joshua.jones@plymouth.ac.uk;
Joshua.N.Jones@uea.ac.uk

Citation:

Jones, J. N., Boulton, S. J., Bennett, G. L., Stokes, M., & Whitworth, M. R. Z. (2021). Temporal variations in landslide distributions following extreme events: Implications for landslide susceptibility modeling. *Journal of Geophysical Research: Earth Surface*, 126, e2021JF006067. <https://doi.org/10.1029/2021JF006067>

Received 6 JAN 2021
Accepted 5 JUL 2021

© 2021. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Temporal Variations in Landslide Distributions Following Extreme Events: Implications for Landslide Susceptibility Modeling

J. N. Jones^{1,2} , S. J. Boulton² , G. L. Bennett³ , M. Stokes² , and M. R. Z. Whitworth⁴ 

¹School of Environmental Sciences, University of East Anglia, Norwich, UK, ²School of Geography, Earth and Environmental Sciences, University of Plymouth, Plymouth, UK, ³College of Life and Environmental Sciences, Geography, University of Exeter, Exeter, UK, ⁴AECOM, Plymouth, UK

Abstract Landslide susceptibility models are fundamental components of landslide risk management strategies. These models typically assume that landslide occurrence is time-independent, even though processes including earthquake preconditioning and landslide path dependency transiently impact landslide occurrence. Understanding the temporal characteristics of landslide occurrence remains limited by a lack of systematic investigation into how landslide distributions vary through time, and how this impacts landslide susceptibility. Here, we apply Kolmogorov-Smirnov and chi-square statistics to a 30-yr inventory of monsoon-triggered landslides from Nepal to systematically quantify how landslide spatial distributions vary through time in “normal” years and years impacted by extreme events. We then develop binary logistic regression (BLR) susceptibility models for 12 yrs in our inventory with >400 landslides and use area under receiver operator curve validation to assess how well these models can hindcast landslide occurrence in other years. Landslide distributions are found to vary through time, particularly in years impacted by storms (1993 and 2002), earthquakes (2015), and floods (2017). Notably, Gorkha earthquake landscape preconditioning shifted 2015 monsoon-triggered landslides to higher slopes, reliefs, and excess topographies. These variations significantly impact BLR susceptibility modeling, with models trained on extreme years unable to consistently hindcast landslide occurrence in other years. However, developing BLR models using increasingly long historical inventories shows that susceptibility models developed using >6–8 yrs of landslide data provide consistently good hindcasting accuracy. Overall, our results challenge time-independent assumptions of landslide susceptibility approaches, highlighting the need for time-dependent modeling techniques or historical inventories for landslide susceptibility modeling.

1. Introduction

Landslides are globally occurring natural hazards posing significant threats to life and sustainable development (Froude & Petley, 2018). Statistics from the Center for Research on the Epidemiology of Disasters reveal that landslides account for 17% of all fatalities due to natural hazards (Sassa & Canuti, 2009), whilst World Bank research suggests >66 million people live in high risk landslide regions (Dilley, 2005). Such statistics highlight a compelling need for efforts to manage and mitigate landslide risk with constant evaluation and improvement.

A fundamental component of landslide management strategies are landslide susceptibility models, which predict the likely geographic locations of future landslides (Guzzetti et al., 2006). Landslide susceptibility models can be developed using physically based techniques (e.g., Goetz et al., 2011; Park et al., 2019; H. J. Wang et al., 2019) or statistical approaches (e.g., Aditian et al., 2018; Baeza & Corominas, 2001; Lee et al., 2008; Reichenbach et al., 2018). Since physically based techniques require vast quantities of empirical data, statistical approaches are often the only viable method to assess landslide susceptibility across regional to global scales, or in data-scarce locations. Most statistical landslide susceptibility approaches utilize the principals of uniformitarianism, with an assumption that, for a given region and trigger type, the spatial distributions of past landslides will be sufficiently similar to those of future landslides so as to facilitate basic prediction (Aleotti & Chowdhury, 1999). Typically, statistical susceptibility models are trained using past landslide data contained within either event, seasonal, or historic inventories (Guzzetti et al., 2012). An event inventory contains landslide data from a single discrete triggering event (e.g., Roback

et al., 2018). A seasonal inventory contains landslides that occurred within a defined time interval, such as a monsoon season (e.g., Fiorucci et al., 2011), and an historic inventory contains all landslides visible across a given region, likely associated with a range of unidentified or undated triggering events (e.g., Jaiswal et al., 2011; Martha et al., 2012). Such multi-temporal seasonal and historic inventories are typically developed using satellite-based monitoring of a given region (e.g., Behling et al., 2016). Landslide susceptibility models are commonly applied across a range of spatial scales (Cascini, 2008), from individual slope units (e.g., Alvioli et al., 2016; Amato et al., 2019) to catchments (e.g., Conforti et al., 2012; Romer & Ferentinou, 2016) to geographical regions (e.g., Sabatakakis et al., 2013; Thi Ngo et al., 2020) and even globally (e.g., L. Lin et al., 2017; Stanley & Kirschbaum, 2017). Statistical landslide susceptibility modeling is very common (Reichenbach et al., 2018), and is often a fundamental component of landslide hazard analyses, risk assessments, land-use planning, and early warning systems (e.g., Fell et al., 2008; Palau et al., 2020; van Westen et al., 2008). However, despite the pervasiveness and importance of such models, as outlined in the following sections there are several fundamental limitations and areas of uncertainty that require further investigation.

First, as implied above, statistical susceptibility approaches assume that, for a given region and trigger event, the likelihood of landslide occurrence is time independent (i.e., landslide locations associated with a given trigger event and region will remain broadly static through time). However, it is well described that climatic, tectonic, and anthropogenic drivers, all of which can influence landslide susceptibility, can vary significantly over long (>1,000 yr) timescales (e.g., Bennett et al., 2016; Kirchner et al., 2001; Larsen & Montgomery, 2012; Molnar & England, 1990; Rahaman et al., 2009; Syvitski & Kettner, 2011), thus invalidating the assumption of long-term landslide time independency (Lombardo et al., 2020). Furthermore, there is now growing evidence to suggest that landslide occurrence cannot be assumed to be stationary over short time periods. For example, landslides in Italy have been shown to exhibit path dependency, whereby the presence of past landslides increases the propensity for new landslides over a period of 10–15 yr (Samia et al., 2017, 2020). Similarly, it is now well established that earthquake induced landscape damage can cause transient increases in the rates of landslide occurrence over annual to decadal timescales (Hovius et al., 2011; Marc et al., 2015, 2019; Parker et al., 2015), a process termed earthquake preconditioning (Parker et al., 2015). Such preconditioning is a likely result of temporary seismic damage that accumulates on positive curvature hillslopes, mountain ridges, and other topographic excesses due to the topographic amplification of seismic ground motion (e.g., Von Specht et al., 2019). There is also evidence to suggest that earthquakes can transiently change landslide spatial distributions, rather than just landslide rates, across short timescales. For example, not only do coseismic landslides tend to lie at higher elevations when compared to rainfall triggered landslides (e.g., Densmore & Hovius, 2000; Rana et al., 2021), but as observed following the 1999 ChiChi event, earthquakes can also transiently shift the locations of post-earthquake rainfall-triggered landslides to higher slope angles (C. W. Lin et al., 2006). Collectively, the above observations challenge the assumption that landslide spatial distributions, and thus landslide susceptibility, are time independent, even over short timescales. However, in the absence of systematic investigations into how landslide spatial distributions change through time, our understanding of the potential impacts of time-dependent landslide occurrence on landslide susceptibility remains limited, thus representing a major topic of uncertainty with important implications for landslide risk modeling.

Second, there remains uncertainty concerning the optimal length of the landslide datasets required to train accurate and reliable susceptibility models. Rarely, if ever, have studies compared the impacts of different inventory lengths (i.e., event vs. seasonal vs. historical) on susceptibility model predictive power. Galli et al. (2008) found that susceptibility models developed using a multi-temporal inventory covering six time slices between 1941 and 1997 outperformed models developed from reconnaissance and geomorphological inventories covering just one and three time slices, respectively. However, in the absence of further investigation, it remains uncertain whether event, seasonal, or historic inventories should be expected to produce the optimal susceptibility model. As multi-seasonal or historic inventories represent the average landslide distributions across multiple years, it could be expected that these would produce more reliable susceptibility models compared to event or single-seasonal inventories, which may be anomalous or unrepresentative of the longer-term average. However, even if true, it would be beneficial to better understand the length of multi-event/seasonal or historical inventory required for reliable and accurate susceptibility modeling.

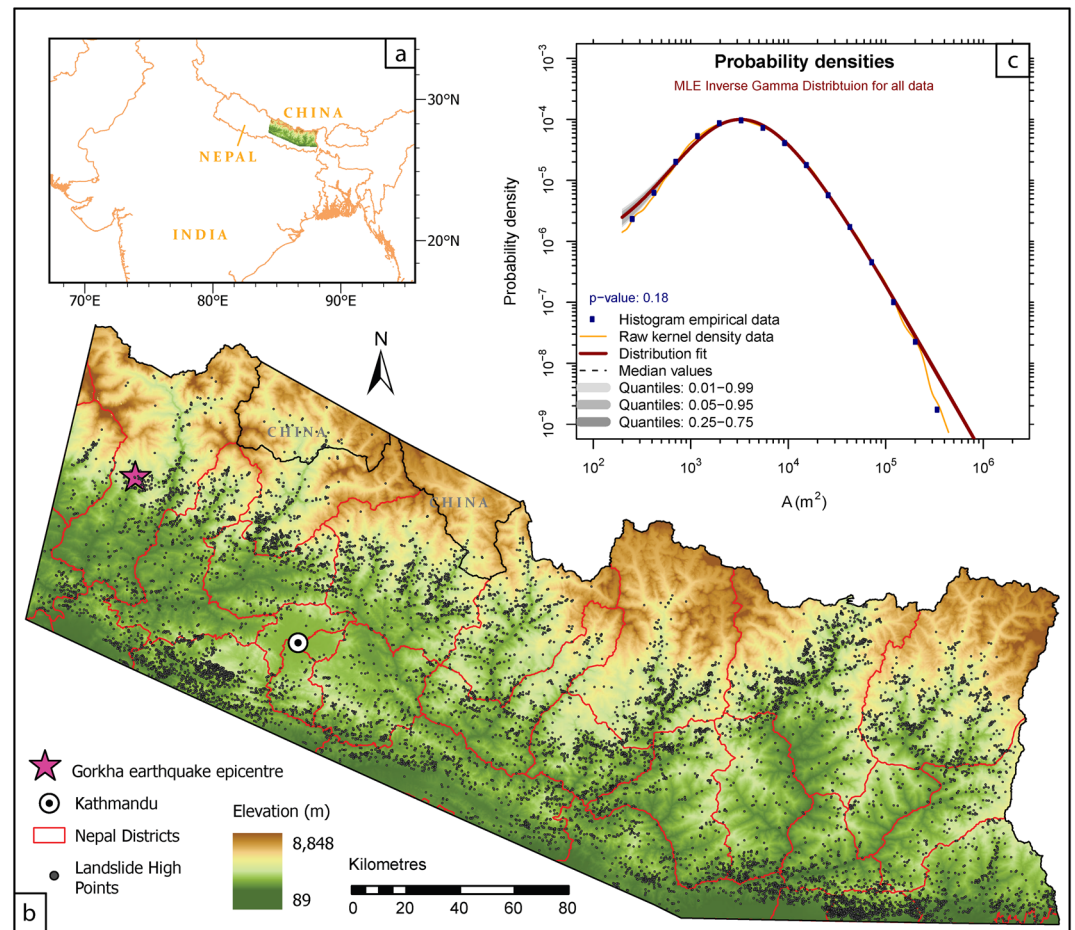


Figure 1. (a) Regional location of our study region. (b) The locations of all 12,838 mapped landslide points used within this study, as well as the locations of Kathmandu and the 2015 Gorkha earthquake epicenter. (c) The inverse-gamma distribution fitted to the probability density function of landslide area for all mapped landslides.

Therefore, this article has three main objectives that seek to address these outstanding issues. First, is to use a newly developed 30-yr multi-seasonal monsoon-triggered landslide inventory for the Nepal Himalaya to systematically investigate whether the spatial distributions of rainfall-triggered landslides vary through time, particularly in response to the 2015 Gorkha earthquake and other “extreme” events. We define an “extreme” event as any high magnitude landslide driver such as, storm, flood, or earthquake preconditioning that triggers landsliding above that expected from the normal monsoon season. Second, is to quantify the impacts of any observed temporal variations in landslide distributions on the hindcasting ability of binary logistic regression (BLR) based susceptibility models. Third, is to utilize our long record of landslide data to investigate how the choice of landslide data used to train BLR susceptibility models influences the model’s hindcast ability. Specifically, we compare the hindcasting ability of models developed using single extreme season inventories versus increasingly long period pseudo-historical (multi-seasonal) landslide inventories, to assess what inventory length is required for consistently accurate temporal modeling.

2. Study Region

The study area is a $\sim 45,000 \text{ km}^2$ region of the central-eastern Nepal Himalaya (Figures 1a and 1b). This region was selected for several reasons. First, Nepal’s dynamic geology, high topography and monsoonal climate make landslides in this region extremely pervasive, with an average of 78 landslide-related fatalities per year (Petley et al., 2007). This fatality rate is estimated to account for $\sim 10\%$ of global rainfall-triggered landslide fatalities (Froude & Petley, 2018), highlighting that from a risk mitigation and management

perspective, Nepal is a region that will directly benefit from an improved understanding of landslide susceptibility. Second, this region has experienced several extreme events over the past three decades, presenting a unique opportunity to investigate how landslide spatial distributions vary in response to a range of different processes. For example, in 2015, the M_w 7.8 Gorkha earthquake struck central-eastern Nepal. This earthquake induced maximum PGAs of 250 cm/s^2 , caused over 8,000 fatalities and reportedly triggered over 240,000 coseismic landslides (Roback et al., 2018; Takai et al., 2016). The Gorkha earthquake also caused enhanced rates of post-earthquake monsoon-triggered landslides (Marc et al., 2019), with high densities of post-earthquake landslides shifting upslope and northward (Kincey et al., 2021). This region has also been impacted by two cloud outburst storms over the past 30 yr, in 1993 (Dhital, 2003) and in 2002 (Dahal & Hasegawa, 2008). On July 19–20, 1993, over a 24 h period, $>540 \text{ mm}$ of rainfall in reportedly fell across a 500 km^2 region of the Kulekhani watershed, 30 km southwest of Kathmandu, causing over 1,500 fatalities (Dhital, 2003), many of which were associated with landslides (Petley et al., 2007). Similarly, on July 23, 2002 in 24 h, $>300 \text{ mm}$ of rainfall reportedly fell across a $14,000 \text{ km}^2$ region of south-central Nepal, causing over 427 fatalities (P. P. Paudel et al., 2003). Furthermore, significant floods impacted several catchments across central-eastern Nepal in 2017 (Gautam & Dong, 2018; Thapa et al., 2020), reportedly destroying and damaging 41,625 and 150,510 houses, respectively, at a cost of at least USD 187.9 million (Government of Nepal, 2017). Finally, as Nepal has an annual monsoon season (June–September), it lends itself to the development of multi-seasonal landslide inventories, where individual monsoon periods can be mapped as discrete seasonal inventories and multiple monsoon seasons can be combined to create pseudo-historical inventories.

3. Data

To investigate temporal variation in landslide spatial distributions and the differences between single-seasonal and historic inventories, a multi-seasonal inventory of rainfall-triggered landslides was developed across the study region (Figures 1a and 1b). Data were also acquired for multiple factors that are commonly observed to control the spatial distributions of rainfall-triggered landslides. Here, we consider 17 possible landslide control factors: elevation, slope, aspect, planform and profile curvature, local relief, excess topography, specific stream power (SSP), channel normalized steepness index (k_{sn}), distance to from landslide crown to channels and major trunk roads, bedrock geology, landuse, Permafrost Index (PFI), average 30-yr May–September rainfall total, annual May–September rainfall and peak monthly May–September rainfall. The following sub-sections describe both the landslide and control factor datasets, including all relevant data collection methodologies.

3.1. Landslide Data

A multi-seasonal inventory of medium to large ($>1,000 \text{ m}^2$) rainfall-triggered landslides covering the period 1988–2018 was manually mapped using Landsat 4/5/8 imagery. In total, 12,838 landslides were mapped within the study region across 29 time slices. The landslides in each time slice were mapped through visual interpretation of pre- and post-time slice Landsat imagery (the specific dates and satellites types used to map each individual landslide feature are shown in the attribute table of supplementary data 1). Each time slice encompassed a single monsoon season (approximately May–September) plus a varying number of non-monsoon months either side. The varying number of non-monsoon months is a result of only using imagery with $<10\%$ cloud cover, which limited the dates of available imagery. As such, the cloud cover was consistent between time slices, but the dates (and possibly snow cover) between the pre- and post-imagery used to map each time slice are not. However, as shown in Figure S1, we found no correlation between the number of non-monsoon months included in a given time slices and the number of landslides included in that times slice, suggesting that this has not unduly impacted the mapping results. Furthermore, landslides were not mapped for 2011 and 2012, since this period was only covered by Landsat 7, which suffered $\sim 22\%$ data loss due to a scan-line failure of the Enhanced Thematic Mapper Plus sensor (Alexandridis et al., 2013).

Each mapped landslide was unique to the time slice in which it was triggered and delineated as a polygon that included both the source and deposition zones. Landslides in the inventory were not differentiated by type, with all types of landslide (flows, slumps, falls, etc.) included, though care was taken to avoid mapping mass-wasting associated with human activity such as cut-and-fill practices and road-tipping. A

basic uncertainty analysis was undertaken whereby all mapped landslides were checked and if necessary edited several months after initially being mapped. As recommended by Marc and Hovius (2015), care was taken to avoid amalgamating multiple separate failures into single polygons. In terms of inventory completeness, following the methods of Malamud et al. (2004), we fit an inverse gamma distribution to the probability density function of landslide area for all mapped landslides using Landsat software (version 10; Rossi & Malamud, 2014; Figure 1c). This shows a rollover of landslide frequency at $\sim 4,000 \text{ m}^2$, which is un-surprising given the resolution of the imagery used to map the landslides, and our focus on medium-to large-scale landslides. The scale of this inventory is such that even though it lacks the smallest landslides, it is well suited for the preliminary modeling of landslide susceptibility, hazard, and risk across regional to national scales (Fell et al., 2008). Although we initially mapped landslide polygons for use beyond the scope of this article, all of the analysis undertaken here uses only the triggering point of each landslide, which we assume to be the highest point along the landslide crest of each mapped polygon (Lombardo & Mai, 2018). There are several reasons why we use points rather than the full polygons. First, BLR susceptibility approaches such as that used here commonly require landslides to be assigned to a single point or cell (e.g., Lombardo & Mai, 2018). Second, much of our analysis requires each landslide to be assigned a single value for each of control factor of interest. This can be done using entire polygons (e.g., Robinson et al., 2017). However, since many of the landslides have large runout zones, we have used the highest point of each landslide crest to represent the triggering conditions rather than an average across the entire landslide polygon. Figure 1b shows the locations of the 12,838 landslide points used within this study.

3.2. Topographical and Hydrological Data

All topographic data were derived from the ALOS World 3D—30 m (AW3D30) version 2.1 global digital surface model (DSM). Elevation, local relief (1,000 m window) slope, aspect, planform curvature, and profile curvature were derived directly from the ALOS DSM using the ArcGIS spatial toolbox. Excess topography, which is a measure of total rock volume above a specified threshold hillslope angle (Blöthe et al., 2015), was extracted from the DSM using the “excess topography” function in the Matlab TopoToolbox using a threshold angle of 30° (Schwanghart & Scherler, 2014).

The ALOS DSM was also used to extract three hydrological factors that are commonly used to model rainfall-triggered landslide occurrence (e.g., Reichenbach et al., 2018). First, the TopoToolbox “STREAMObj” function was used to extract the stream channel network across the study region for a threshold upstream area of 1 km^2 . Euclidean distances to these channels, with a 30 m buffer, were then extracted for every landslide and cell within the study region using standard ArcGIS distance tools. Second, the normalized steepness index (k_{sn}) for those channels was extracted using the TopoToolbox “ k_{sn} ” function with a concavity (θ) value of 0.36 as determined by the TopoToolbox “slope area” function. Finally, the SSP of those channels, which is total stream power per unit channel width, was calculated using Equation 1:

$$\Omega = \frac{\rho g Q S}{w} \quad (1)$$

where p is the density of water ($1,000 \text{ kg/m}^3$), g is acceleration due to gravity (9.81 m/s^2), S is the energy gradient, or channel slope (derived from the DSM using standard ArcGIS tools in units of m/m), Q is channel discharge (derived from the DSM using standard hydrological ArcGIS tools in units of m^3/s), and W is channel width (calculated as a function of discharge according to the scaling relationships of Craddock et al. (2007) in units of m).

3.3. Geological Data

Like many developing countries with extreme, highly inaccessible terrain, Nepal lacks high-resolution geological data. The Nepal Department of Mines and Geology hold 1:250,000 regional scale maps of the main lithologies mapped across the country. These maps were digitized and then simplified into the following groups: dolomites, granites and gneisses, phyllites, Quaternary sandstones and conglomerates, marbles, schists, quartzites, shales, and “undifferentiated”.

3.4. Landuse and Road Data

Since Nepal lacks freely available field-validated landuse data, we use the landuse classifications of the ESA-GlobCover product. This product defines 22 land cover classes (Bontemps et al., 2011) and has an estimated accuracy of 73% (Defourny et al., 2009). For simplification, we reclassify these 22 classes into the following: cropland, forest, shrub and grassland, bare/sparse earth, water, permanent snow/ice, and artificial. The main limitations of this data set are its low (300 m) spatial resolution and the fact that it is based on satellite data from 2009 only (see Section 6.1). As such, to assess the validity of the ESA data, we undertook additional sensitivity analysis (see Section 6.1) using the MODIS Land Cover (MCD12Q1) dataset. This dataset is a global landuse product, with a 500 m spatial resolution and annual time step from 2001—present, that uses 13 different classification schemes (Sulla-Menashe & Friedl, 2018).

Road data across the study region were obtained from the Open Street Map (Humanitarian Data Exchange, 2020). All primary, secondary and trunk roads were extracted from this dataset, and the Euclidean distances to these roads, with a 30 m buffer, was calculated using standard ArcGIS distance tools for every landside and cell within the study region.

3.5. Climate Data

As permafrost is a key feature of high mountain environments it is necessary to consider its influence on landslide occurrence. Permafrost is defined as sub-surface material with a temperature of $\leq 0^{\circ}\text{C}$ for at least two consecutive years (French et al., 1988). Here, we use a global dataset of permafrost extent developed by Gruber (2012). This PFI classifies cells with values between 0.01 and 1, where a value of 1 suggests that permafrost will be present under all conditions and a value of 0.01 suggests that permafrost will be present under ideal meteorological conditions only. This dataset has a spatial cell resolution of 500×500 m.

Since this paper focuses on rainfall-triggered landslides, it is necessary to consider rainfall covariates. In Nepal, it is estimated that $>90\%$ of rainfall-triggered landslides occur during the annual monsoon season (May–September; Petley et al., 2007; Stanley et al., 2020) when Nepal experiences $>80\%$ of its yearly rainfall (Dahal & Hasegawa, 2008). There are a number of products that provide precipitation data across the study region (e.g., Ashouri et al., 2015; Beck et al., 2017; Sun et al., 2018; X. Wang et al., 2021; Yatagai et al., 2012). However, for this study, we specifically required precipitation data extending back to 1988 with a spatial resolution of at least $0.25^{\circ} \times 0.25^{\circ}$. One of the few precipitation products to meet these specifications is the PERSIANN-CDR Climate Data Record (Sun et al., 2018). This product covers the period 1983 to present at a spatial resolution of $0.25^{\circ} \times 0.25^{\circ}$ and a daily to monthly temporal resolution (Ashouri et al., 2015). Here, we derive three rainfall metrics based on this data: the mean 30-yr total precipitation for the months May–September, annual May–September total precipitation, and annual peak monthly May–September precipitation.

4. Methods

The objectives of this paper are to quantify how the spatial distributions and susceptibility of rainfall triggered landslides varies through time, and to assess the landslide data required to develop robust and reliable susceptibility models. We did this using several methods as outlined in Figures 2a–2c. First, to assess how the spatial distributions of landslides vary through time, we use chi-square and two-sample K-S (Kolmogorov-Smirnoff) analysis to compare the distributions of landslides with respect to several control factors (see x-axis labels of Figure 3) across every mapped time slice of our inventory. Second, to assess how any temporal variations in landslide distributions impact a commonly used type of landslide susceptibility approach, we used BLR, implemented alongside a least absolute shrinkage and selection operator (LASSO) for covariate selection, to develop susceptibility models for 12 separate time-slices of our inventory. We then used area under receiver operator curve (AUROC) validation to assess how well the developed BLR susceptibility models for 1 yr could predict the landslide data for other years, and whether BLR models developed with increasingly longer period pseudo-historical inventories have increasing hindcasting power relative to models developed from single extreme seasons. The following Sections 4.1–4.3 and Figures 2a–2c outline each of these methodologies.

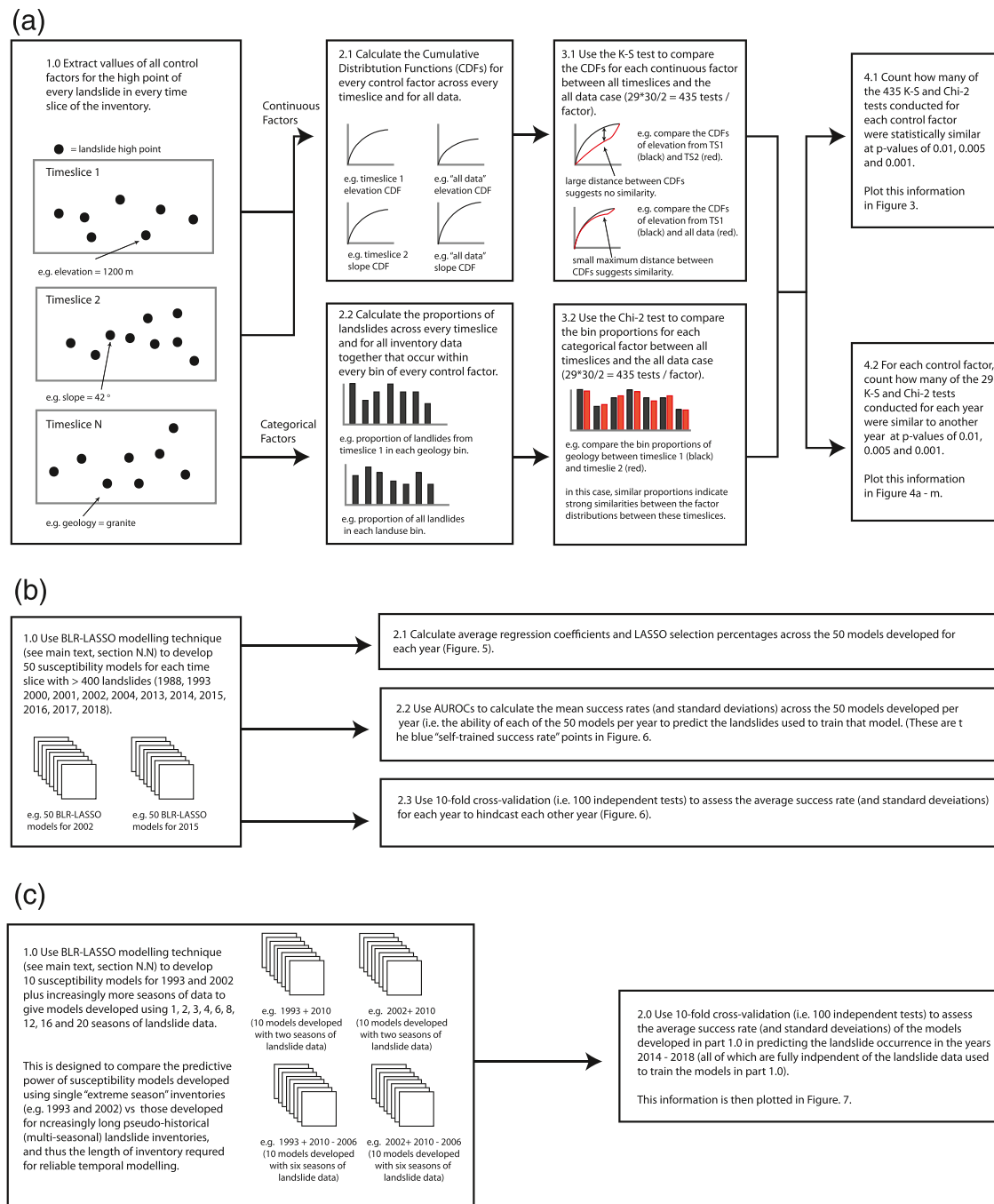


Figure 2. Workflow methodology for (a) comparing landslide distributions in different time slices using the K-S and chi-square statistics. (b) Developing and validating landslide susceptibility models using the BLR-LASSO approach. (c) Developing and validating landslide susceptibility models using the BLR-LASSO approach for increasingly long pseudo-historical (multi-seasonal) landslide inventories.

4.1. K-S and Chi-Square Analysis

The two sample K-S test compares the similarity between two continuous samples, operating under the null hypothesis that the two samples being compared are similar. The null hypothesis can be rejected if the p -value of the K-S statistic, which measures the largest distance between the empirical cumulative distribution functions (CDFs) fitted to the continuous samples being compared (e.g., Figure 2a), is lower than a given α value. Similarly, the chi-square test evaluates whether two categorical samples have similar underlying

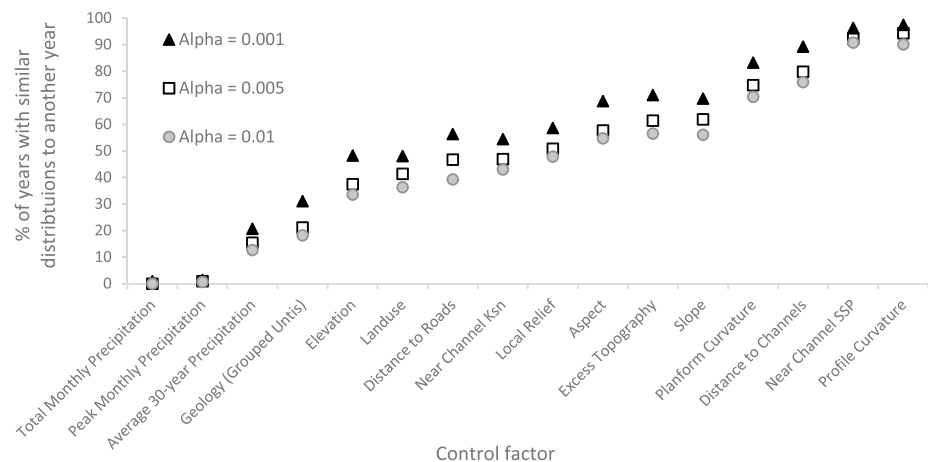


Figure 3. For each control factor, the percentage of the 435 K-S or chi-square tests that showed similarity between 2 yr at different alpha values.

distributions, operating under the null-hypothesis that the two categorical samples have similar proportions of data in each category bin (e.g., a land-use class or geological unit). The K-S and chi-square tests were applied as follows to assess how the spatial distributions of the landslides in our inventory vary through time:

1. For every landslide, the values of all continuous and categorical control factors of interest were extracted at the highest elevation point (the assumed triggering location) of each landslide.
2. For the continuous control factors, the empirical CDFs were calculated for the landslides in every individual time slice, and for all of the time slices combined (the “all data” case). Similarly, for the categorical control factors, the proportion of landslides in each control factor bin were calculated for the landslides in every individual time slice, and for all of the time slices combined (the “all data” case).
3. For the continuous control factors, the K-S test was used to compare the CDFs of every time slice to (1) the CDFs of every other time slice and (2) the CDF of the “all data” case. With 29 time slices plus the “all data” case, this gave 435 individual comparisons ($30 \times 29/2$). Similarly, for the categorical control factors, the chi-square test was used to compare the proportions of landslides per bin in every time slice to the proportions in every other time slice plus the proportions of the “all data” case. Again, this gave 435 individual comparisons.
4. For each continuous and categorical control factor, the number of K-S or chi-square tests with p -values below significance (alpha) thresholds of 0.01, 0.005, and 0.001 were counted. From this, the percentage of time slices with statistically similar landslide distributions to another time slice or the “all data” case were obtained (Figure 2a).
5. The results obtained from (4) give an overall appreciation of how similar the landslide distributions with respect to each control factor are through time. However, this does not show where any dissimilarities are temporally concentrated. For example, if a given control factor is found to be similar in only 50% of the 435 tests, it is not known whether this is because every year was only similar to 50% of the other years, or because 50% of the years were similar to no other years. This information is important for identifying where in time any dissimilarities actually occurred. So, for each control factor, we also counted how many of the other time slices and the “all data” case each individual time slice was similar to (Figure 2a).

4.2. Binary Logistic Regression Modeling

The K-S and chi-square tests allow quantification of the temporal variation in single landslide control factor distributions, considered independently from any others. The second objective of this paper was to assess how typical susceptibility models developed for different time slices of our inventory using the control factors investigated previously might differ. Here, we used BLR based susceptibility modeling implemented alongside a LASSO (e.g., Lombardo & Mai, 2018) for variable selection using the glmnet package (Hastie & Qian, 2014) in the statistical software, R. The reason for using logistic regression is that this approach

has historically been, and remains, by the far most common method of assessing landslide susceptibility (Reichenbach et al., 2018), so it is particularly vital to understand how temporal variance affects this model type. However, it should be noted that as this paper only uses logistic regression, it is possible that other susceptibility techniques such as machine learning and fuzzy logic (e.g., Kritikos et al., 2015), which can outperform logistic regression in some cases (e.g., Braun et al., 2015), will not be similarly affected by the issues of temporal variance investigated here.

The BLR approach to landslide susceptibility modeling is common in the literature (e.g., Reichenbach et al., 2018) so the mathematics behind BLR modeling are well described elsewhere (e.g., Appendix A of Lombardo & Mai, 2018) and will not be re-described here. The implemented LASSO is based on an algorithm that uses cyclical coordinate descent (Friedman et al., 2010) to optimize the objective function using 10-fold cross-validated penalized maximum likelihoods (Hastie & Qian, 2014). In other words, it cycles through multiple different combinations of control factor variables by systematically setting some factors to zero, until it converges on the optimum solution (Friedman et al., 2010). In these models, the objective function (or solution) to be optimized is the self-validated model success rate obtained from the AUROC (see Section 4.3) of each model, that is, the success of each model at classifying the data used to train that model. The advantage of this methodology is that as well as providing results on control factor regression coefficients, you also obtain results on control factor selection percentage. This is particularly useful in this case as it allows a quantitative assessment of whether the optimal control factors for landslide susceptibility assessment are stable through time. This approach allows the optimum diagnostic performance (or hind-casting ability) of our different models to be compared. However, having different sets of control factors between models can influence the individual regression coefficients assigned to each control factor, so this limitation requires consideration when interpreting our regression coefficient results.

To ensure the robustness of our results, we opted to run the BLR-LASSO models only for the 12 yr that observed >400 landslides (1988, 1993, 2000–2002, 2004, and 2013–2018). Before running the models, our landslide and control factor data required further processing. First, we divided our study region into a 30×30 m grid of $\sim 5 \times 10^7$ cells. For each year, each grid cell was assigned a value of one if it included a landslide triggering point (landslide presence) and a value of zero if not (landslide absence). All control factor datasets were then resampled to this grid. Then, for each year, we extracted 50 balanced sub-datasets, where each subset included all control factor data associated with that year's landslide presence cells plus an equal number of randomly selected landslide absence cells. In total, this gave us 600 balanced datasets across the 12 yr to be modeled.

Each of these 600 datasets included information on landslide presence or absence, plus the associated values of 17 control factors. Of these 17 factors, two were categorical (geology and landuse) and the rest were continuous. To ensure the final regression coefficients calculated for each factor were objectively comparable, we rescaled the continuous factors using zero-mean unit variance (e.g., Lombardo & Mai, 2018). Furthermore, with so many factors, it was possible that some would be collinear. This is potentially problematic, as significant collinearity between factors can introduce error and instability into regression models (Zuur et al., 2010). As such, before inputting our datasets into the glmnet model, we tested for collinearity between all factors in all 600 sub-datasets using the Variance Inflation Factor (VIF) functions of Zuur et al. (2010). The VIF is a common measure of multi-collinearity in a set of regression variables, that is, equal to the ratio of the variance in a multi-variable model to the variance of a model that only includes a single independent variable. VIF's can be calculated for each individual variable, and a $VIF > 5$ suggests that the associated independent variable is highly collinear with at least one other variable in the model. Here, for each dataset, we calculated the VIF's for each independent control factor. Then, if any factors had VIF's > 5 , the factor with the highest VIF was removed and the VIF's were recalculated for the remaining factors. This was repeated until all factors had VIF's < 5 . In this case, the "total rainfall" factor was found to be highly collinear with peak rainfall, average rainfall and elevation. Once total rainfall was removed, all other factors had VIF's of < 5 . Finally, before running the glmnet model, our two categorical factors were coerced into dummy variables (i.e., presence/absence for each sub-category).

The processed 50 balanced subsets for each of the 12 modeled years were then run through the glmnet model, where each model used the 10-fold cross-validated LASSO for factor selection. The resulting factor selections and associated regression coefficients were then averaged for each year based on all 50 models for

that year. This allowed quantification of how each factors selection percentage and regression-coefficient varied through time across each of the 12 modeled years.

4.3. AUROC Model Validation

The second and third objectives of this article are to investigate the influences of temporal variations in landslide distributions and of the choice of landslide inventory on susceptibility model predictive power. To assess how the temporal variations in landslide distributions quantified here influence overall susceptibility predictability, we use AUROC validation to assess how well the developed BLR models for each year could predict the landslide data from each of the other modeled years (i.e., each of the other years: 1988, 1993, 2000–2002, 2004, and 2013–2018). The receiver operating curve (ROC) is a probability curve obtained by plotting the true positive rate (TPR) against the false positive rate (FPR). The TPR, which is commonly referred to as sensitivity, is in this case the proportion of landslide presences that were correctly classified as landslide presences. The FPR, which is commonly referred to as 1—specificity, is the proportion of landslide absences that were incorrectly classified as landslide presences. The AUROC value indicates the degree to which a binary model correctly predicted the observed classes, in this case landslide presences and absences. An AUROC value of one indicates that a model was 100% accurate, whilst an AUROC value of 0.5 is equivalent to the result of a random predictor with no classification capacity. A value <0.5 indicates that a model is actively reciprocating the classification, that is, in this case would be predicting landslide absences as presences and vice versa. Typically, a models with AUROC values of 0.7–0.75 are considered good, whilst models with AUROC values >0.8 are considered very good (e.g., Marjanović, 2013; Vakhshoori & Zare, 2018). Here, we used 10-fold cross validation, whereby 100 AUROC tests between the models developed for a given year and random balanced landslide presence/absence datasets from another year, were used to calculate the average AUROC values and standard deviations for each year's ability to hindcast the landslide occurrence in each other modeled year. If all 12 years had average AUROC values when hindcasting all other years of >0.7–0.75, then it could be concluded that any temporal landslide control factor variation does not actually influence the overall predictive power of landslide susceptibility models developed on a given year. However, if some or all years are incapable of sufficiently hindcasting all or some other years, it would suggest that great care must be taken when developing and applying susceptibility models through time.

Finally, we investigated how the amount of landslide data used to train a susceptibility model influences a models hindcast power. To do this, the following methodology (Figure 1c) was used:

1. Take two single seasons of data (1993 and 2002) that when used in isolation as in Section 4.2, produced susceptibility models that were poor at hindcasting the other years (see Section 5, Figure 6).
2. Systematically add additional seasons of landslide data to both the 1993 and 2002 datasets to create increasingly long pseudo-historical (multi-seasonal) datasets that include 2, 3, 4, 6, 8, 12, 16, and 20 seasons of landslide data. These inventories were developed by adding seasons from 2010 backwards. Thus the “2-season” inventories were developed using the data from 1993 + 2010 and 2002 + 2010, and the “3-season” inventories were developed using 1993 + 2010 + 2009 and 2002 + 2010 + 2009, etc.
3. Use the same methodology as described in Section 4.2 to develop susceptibility models from each of the increasingly long pseudo-historical datasets.
4. Use 10-fold cross validated AUROC analysis as described above to assess how well each of these increasingly long multi-seasonal inventories could predict the landslide data from each of the years 2014–2018.

5. Results

5.1. K-S and Chi-Square Analysis

The K-S and chi-square analysis shows that there is significant variation between the landslide distributions resulting from different control factors in different years (Figure 3). Of the 16 control factor distributions compared across all years (435 tests), only four (profile curvature, near channel SSP, distance to channels, and planform curvature; Figure 3) had >70% similarity across all tests. Factors commonly used to model landslide occurrence such as hillslope gradient, aspect, relief, landuse, elevation, and geology had similar distributions in just 30%–70% of all tests. The least similar factor distributions were average 30-yr

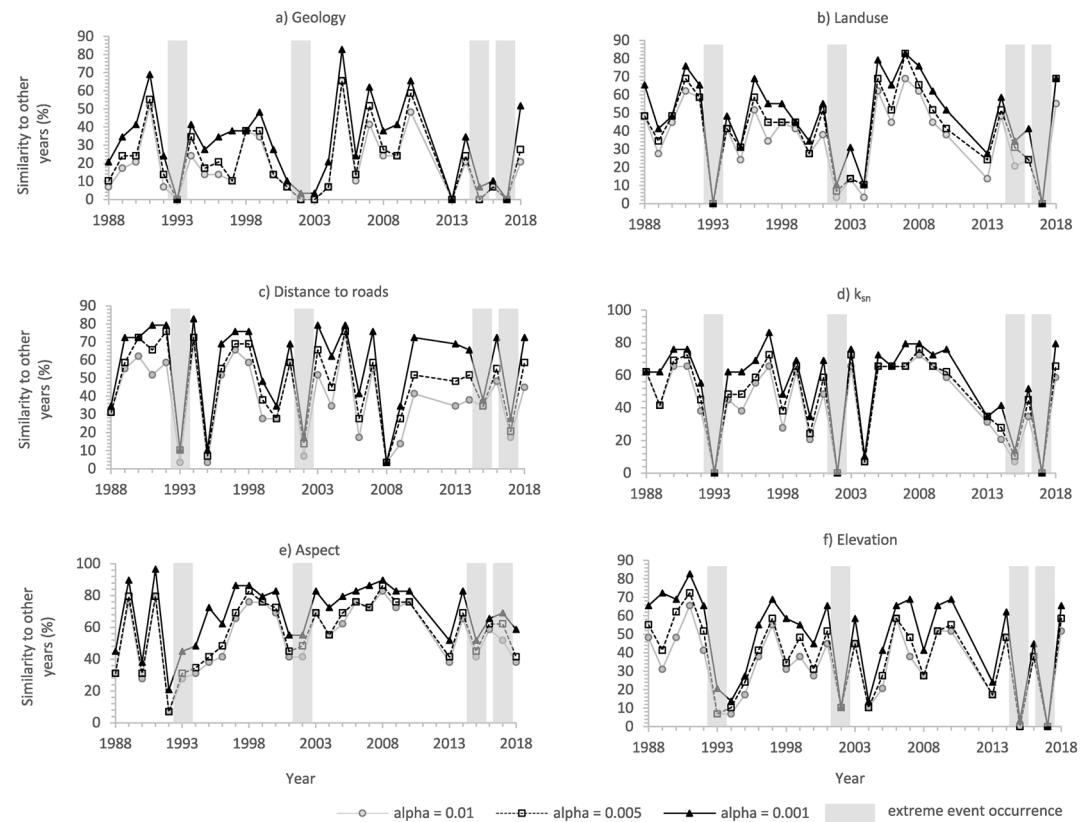


Figure 4. (a–m) For each control factor from Figure 3 with >20% overall similarity, the percentage of the 29 tests per year that showed similarity to at least one other year at different alpha values.

precipitation, peak monthly precipitation and total monthly precipitation, which had maximum similarities across all tests of >20%. However, these low values may be affected by the low (~30 km) spatial resolution of the PERSIANN-CDR data used to obtain these control factor values.

The percentages quoted above are useful for obtaining an overall appreciation of control factor distribution similarity; however, important information about where the dissimilarities are concentrated is hidden. For example, in the case of slope, are the 30%–44% of tests that showed no similarity to any other tests equally distributed across all 29 yr? Or are they due to a small number of years that were totally dissimilar to every other year? To investigate this issue, we calculated for each control factor the percentage of tests by year that were similar to another year (Figures 4a–4m). This reveals that the dissimilarities in some factor distributions (geology, landuse, distance to roads, near channel k_{sn} , aspect, and elevation; Figures 4a–4f) are stochastically spread across multiple years. Conversely, the four factor distributions with the highest overall similarity between years (profile curvature, planform curvature, SSP, and distance to channels; Figures 4g–4j) are relatively stable through time, with only minor dissimilarities in some years. Finally, some covariate distributions (slope, relief, and excess topography; Figures 4k–4m) show a degree of stability across most years, but have a small number of years (notably 1993, 2002, 2015, and 2017) showing significant dissimilarity to all others.

5.2. BLR Modeling

The first outputs of our BLR modeling were the average regression coefficients and LASSO selection percentages obtained for each control factor across the 50 models ran for each year (Figures 5a–5p). These outputs highlight significant differences between the different control factors. The most consistent continuous factors in terms of selection percentage were elevation, slope and PFI (Figures 5a–5c), which were selected in almost 100% of each year's 50 models. The regression coefficients for these factors were also stable in

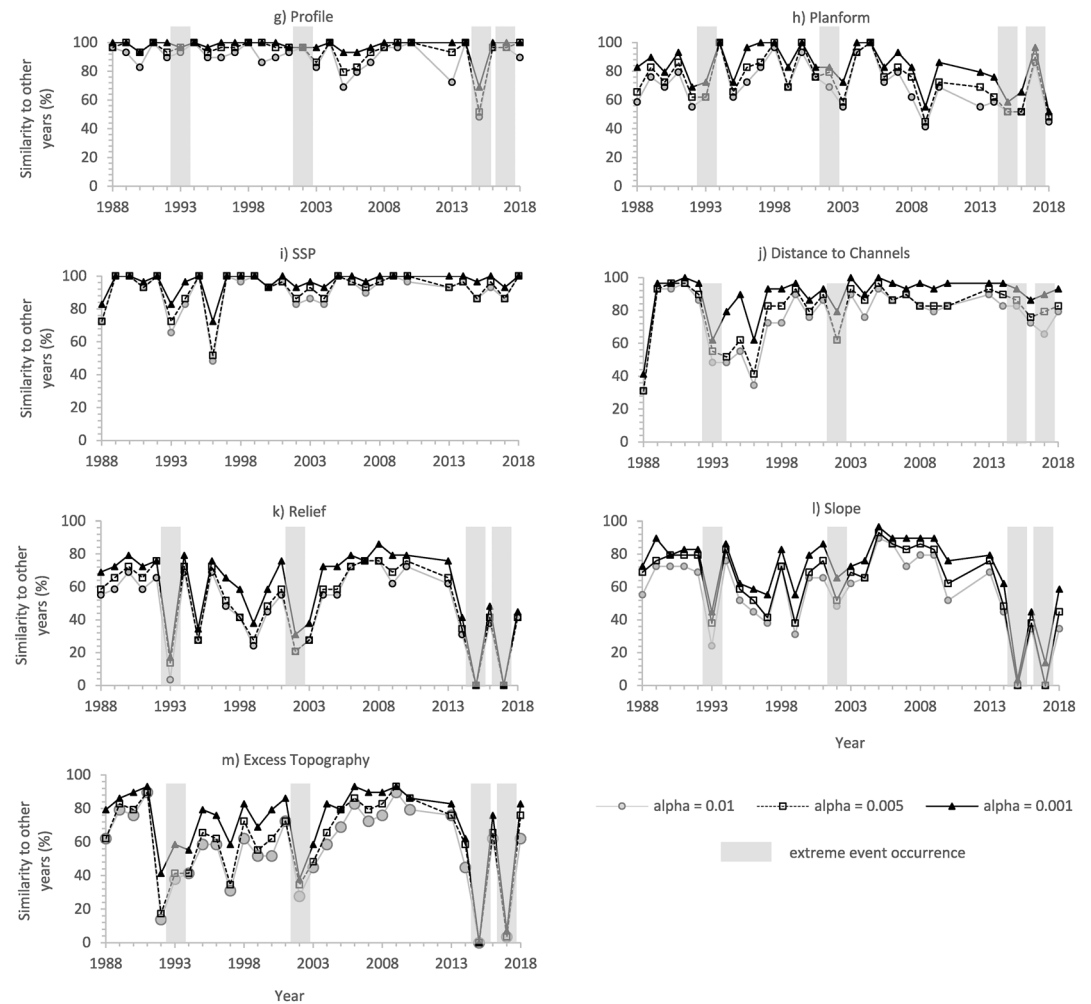


Figure 4. *Continued.*

that they were always all positive (e.g., slope), suggesting that a factor was increasing the probability of landsliding, or negative (e.g., elevation and PFI), suggesting that a factor was decreasing the probability of landsliding. However, in the case of PFI, the 2015 regression coefficient was of notably greater magnitude than the other years, with a value close to -1.0 , compared to values of -0.2 to -0.6 for most other years.

The next most consistent continuous factors in terms of selection percentage were profile curvature, planform curvature and local relief (Figures 5d–5f). Profile curvature was selected by $>80\%$ of models for most years and had regression coefficients of -0.05 to -0.15 in all years but 2000 and 2015, which had slightly greater magnitude coefficients of ~ -0.25 . Planform curvature was less consistently selected than profile curvature, particularly in the year's preceding 2014. The regression coefficients for planform curvature were also more variable, with 1993 having a negative coefficient (-0.05) whilst all other years were positive. Local relief was always selected by $>40\%$ of models for a given year, with near 100% selection in 1993, 2000–2002, and 2015–2016. The regression coefficients for local relief were also highly stable through time, with most years having coefficients of 0.1 – 0.3 . The exception to this was 2015, which had a significantly higher local relief coefficient of 0.8 .

All of the other continuous factors (Figures 5g–5n) were highly variable through time, in terms of both selection percentage and regression coefficients, suggesting that these factors are less reliable in defining landslide spatial distributions. The most notable observation from these factors is that excess topography (Figure 5k), which had an almost 0% selection rate in most years, had selection rates of 30% and 70% in 1993 and 2015 respectively. However, despite 1993 and 2015 both having higher percentage excess topography

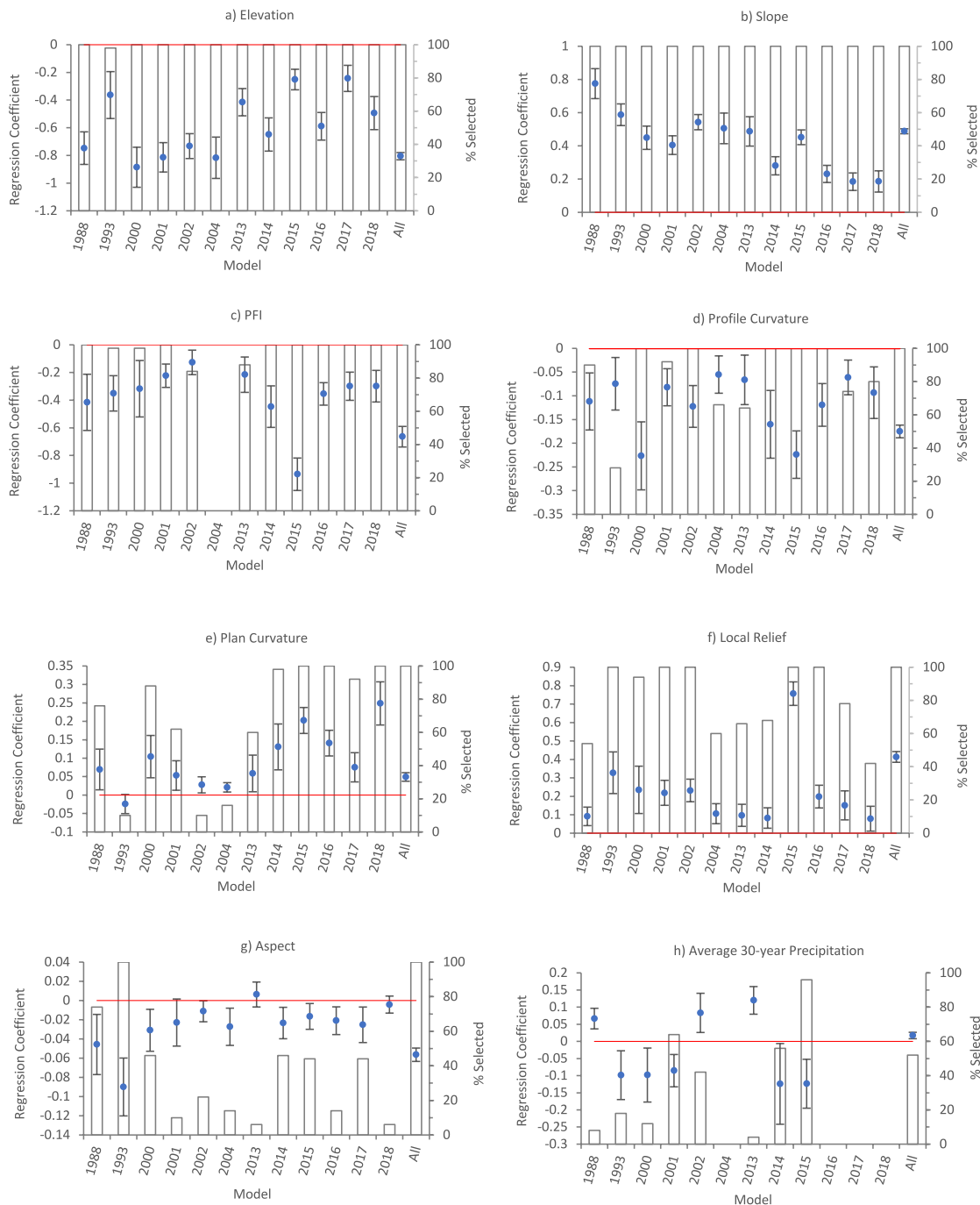


Figure 5. (a–n) Results of binary logistic regression modeling for each control factor. Blue-circles show the average regression coefficient calculated from the 50-models run per year. Error bars show ± 1 SD. Bars show the percentage of the 50 models run for each year in which that control factor was selected by the LASSO. Red line shows the 0-line of the regression coefficient axis. Positive coefficients indicate that a factor is making landslides more likely, whilst a negative coefficients indicate that a factor is making landslides less likely.

selections than other years, their regression coefficients were of opposite sign, with 1993 having a value of -0.1 , and 2015 a value of $+0.1$.

The two categorical factors also show significant variation through time in terms of variable selection and regression coefficients. In terms of geology, the unit most consistently selected by the LASSO is Quaternary sandstone/conglomerate (Figure 5o), which is selected at least 40% of the time in 8 of the 12 modeled years,

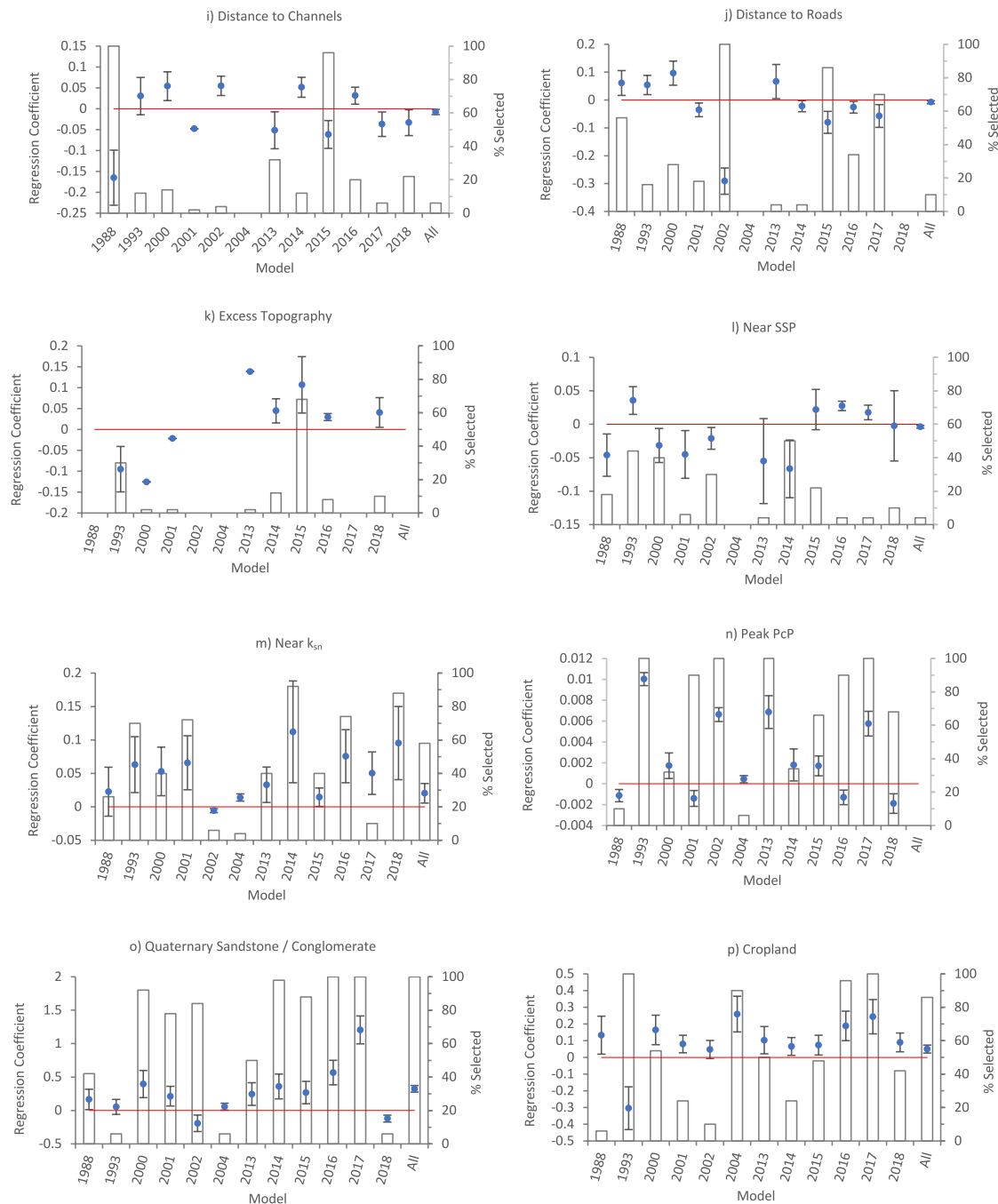


Figure 5. Continued.

and had regression coefficients that were positive for all years except 2002 and 2018. Most other units were rarely selected >50% of the time for a given year and had regression coefficients that fluctuate between being positive and negative (see Supplementary Information; Figures S2a–S2h). In terms of landuse, only cropland (Figure 5p) was consistently selected by the LASSO, with 5 years having >50% selection, and 3 years with >40% selection. Furthermore, the regression coefficients for cropland are stable, with values of 0.05–0.3 for all years except 1993, which anomalously had a value of -0.3 . Of the other landuse variables (see Figures S2i–S2n), permanent snow/ice cover and forest were the next two most selected factors, with both having on average negative coefficients. For permanent snow/ice, as observed with PFI, 2015 differed

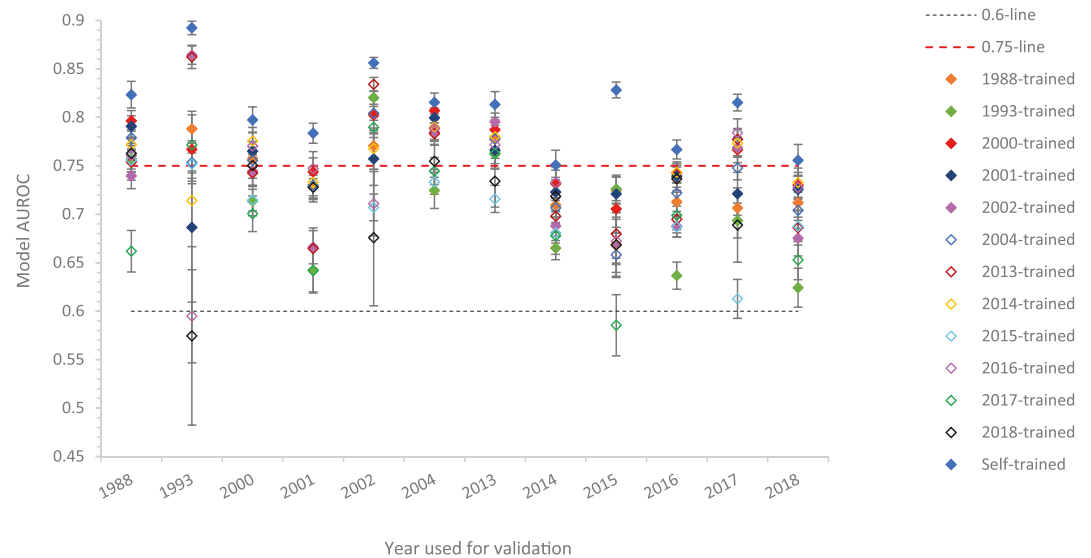


Figure 6. Area under receiver operator curve (AUROC) validation results. This quantifies how accurately the BLR susceptibility models trained on 1 yr landslide and control factor data could predict the landslide occurrences from another year. All results were obtained via 10-fold cross validation, whereby 10-models trained on 1 yr were used to predict 10 random subsets of data from another year, thus giving 100 results per validation from which the averages and standard deviations on this plot were calculated.

from the other years in that it had a higher magnitude coefficient of -1.5 compared to >-0.5 in other years. Shrub/grassland, bare/sparse earth, water, and artificial land showed near-zero selection across all years.

5.3. AUROC Validation

To assess how well each year could hindcast each other year, 10-fold cross-validation, whereby 100 independent AUROC tests between models from 1 yr and data from another year, was used to obtain each reported AUROC value and associated error (Figure 6). We find that there is significant variability in the hindcasting power of models trained on different years. The years 2000, 2001, 2002, 2004, 2013, and 2014 were all reasonably well hindcast by the models developed on other years relative to their self-validated AUROC success rates. In all these years, the years least capable of hindcasting them were typically 1993, 2002, 2015, or 2017. The year 2015 was the hardest for the other years to predict relative to its self-validated AUROC success rate, with no year successfully predicting it with $\text{AUROC} > 0.75$, and 2017 failing to predict it with $\text{AUROC} > 0.6$. The years 2016, 2017, and 2018 were mostly well hindcast by other models relative to their self-validated AUROC values, though again, the 1993, 2002, 2015, 2017, and 2018 models were consistently the least capable of hindcasting them. Overall, these results highlight that some years (1993 and 2015) were particularly hard to hindcast by some models, and that models developed from some years (particularly 1993, 2002, 2015, and 2017) were consistently poor at hindcasting the other years.

Figures 7a–7j shows the AUROC results of hindcasting the years 2014–2018 using susceptibility models developed using the years 2002 and 1993 alone (x-axis value of 1) plus increasingly more seasons of landslide data. In all but one to two cases, the AUROC values fit a positive logarithmic curve (R^2 values of 0.76–0.97), whereby the obtained AUROC values increase rapidly from using 1993 or 2002 alone to using 1993 and 2002 alongside six to eight other seasons of data, with a stabilizing of the AUROC value when using more than approximately eight seasons. This highlights that single seasons with poor hindcasting ability should be combined with at least six to eight other seasons to ensure optimum hindcast accuracy. The main exceptions to this relates to hindcasting the 2017 season using the 2002 data plus other seasons (Figure 7d). Here, the data fit a strong negative logarithmic curve ($R^2 = 0.91$), showing that adding more seasons of data to 2002 was reducing the hindcasting accuracy in this case. Finally, in the case where 1993 was used to hindcast 2015 (Figure 7g), there is still a weak positive logarithmic correlation between number of seasons used and AUROC, though with more variance.

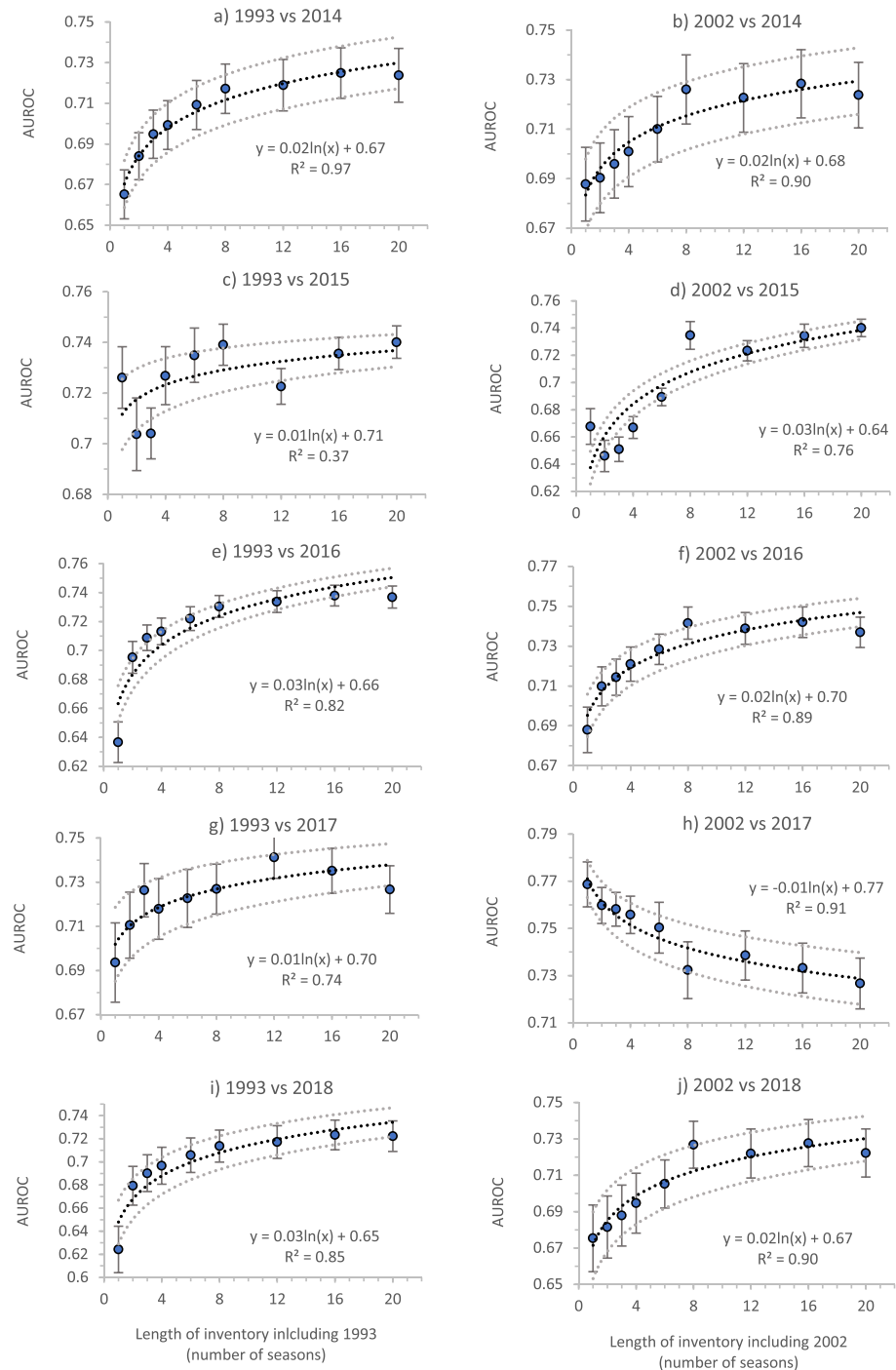


Figure 7. AUROC values obtained from predicting (or hindcasting) the landslide occurrence in the years 2014 (a–b), 2015 (c–d), 2016 (e–f), 2017 (g–h), and 2018 (i–j) using models developed using 1993 and 2002 plus increasingly more seasons of landslide data (i.e., increasingly long pseudo-historical landslide inventories).

6. Discussion

6.1. Data Limitations and Assumptions

Our results show that: (a) There is variation in landslide control factor distributions through time. (b) There is variation in landslide susceptibility results through time. (c) Susceptibility results improve when models

are developed using several seasons of landslide data. However, before discussing the implications and causes of these results in terms of physical processes, it is important to consider whether any of our observations could be due to limitations of the various control factor datasets we have used. The main potential data limitation is that some control factor datasets, notably distance to roads, PFI and land use, are considered as static factors in our modeling, despite the fact that these factors may themselves vary through time. This is an unavoidable consequence of the data-scarce nature of the study region, but does raise the question of whether any of the observed temporal variance in our results is owing to this data limitation, rather than true physical processes. The following sections thus critically appraise the issues relating to the distance to road, PFI and land-use data, highlighting these for the reader and outlining why we do not believe they unduly impact the overall results.

In terms of distance to major roads, road locations will have changed throughout the 30-yr period considered here, with road building initiatives increasing road density, particularly the density of small informal rural roads, across Nepal (Mcadoo et al., 2018). However, as outlined in Section 3.4, in the absence of available annual road data, our “distance to major roads” factor was based solely on the positions of large trunk, primary, and secondary roads in 2016. We only used these larger roads as these types of infrastructure were more likely to have existed for the entirety of our mapped period. For example, construction of the Araniko highway trunk road began in 1961 (Ao et al., 2020; Murton, 2017). However, it is still possible that some of the primary and secondary roads included in our dataset did not exist in the early part of our time period. As such, our distance to roads metric can only be confidently considered as a topographic metric of “distance to road position in 2016.” This factor may not be geomorphologically useful, but our approach with the LASSO is designed to deal with this uncertainty. Indeed, we see that very few models ever select this as an important factor. The main issue with our road data is how it is interpreted. In Figure 4c, the years 1993, 1995, 2002, and 2008 differ most significantly from all other years. Likewise, in Figure 5j, it is only 2002 that has both a high LASSO selection and significantly different regression coefficient to other years. These years can only be confidently interpreted as having different landslide positions relative to roads in 2016 compared to other years, meaning that these variations are not necessarily due to road occurrence, but rather any process that could shift landslide topographic distributions. This is an important distinction that requires acknowledgment when interpreting the results and discussion presented here, but it is not an issue that affects the validity of our overall conclusions.

For the PFI data, to assess the impact of any temporal changes in permafrost on the results, it is first necessary to consider how much the limits of Himalaya permafrost have changed over the mapped period. It is estimated that the lower limit of permafrost (LLP) extent shifted no more than 100–300 m between 1973 and 1991, before remaining relatively stable until at least 2004 (Fukui et al., 2007). Furthermore, the current rate of change of permafrost extent is estimated to be ~ 1.3 – 2.6 m/yr, with a maximum expected potential increase in the LLP of 188 m between 2009 and 2039 (Chauhan & Thakuri, 2017). These studies show that whilst permafrost extent is likely changing in response to climate change, the rate of change is small relative to the $500 \text{ m} \times 500 \text{ m}$ resolution of the PFI data used here, with even the maximum expected changes for the next 20 yr below the resolution of the data set. Furthermore, the permafrost data used here is an index which estimates the likelihood of permafrost based on climatic conditions, where a value of 1 suggests that permafrost will be present under all conditions and a value of 0.01 suggests that permafrost will be present under ideal meteorological conditions only. As a landslide was only classified as being in permafrost if it had a value of 1, our permafrost classifications should not be impacted by small annual changes in permafrost extent as these are most likely to affect index values < 1 . Figure 5c also shows that PFI was consistently selected as an important factor by the LASSO, with all years but 2015 having very consistent regression coefficient values. There are two possible explanations for why 2015 differs. One, is that 2015 had very different permafrost conditions to that estimated by the PFI data used here, and if more accurate PFI data been used, then the 2015 perturbation would not be observed. Or two, as described in Sections 5.2 and 6.2, it is due to a “true” shifting of the 2015 landslide distributions that made the negative relationship between PFI and landslide occurrence even stronger. Literature investigations reveal no evidence to support explanation 1. However, as described in Sections 5 and 6.2, it is known that the 2015 monsoon-triggered landslides were impacted by earthquake preconditioning that shifted their locations relative to other years, supporting explanation 2. Therefore, it is possible that our PFI data may not fully account for changes in permafrost extent and this should be considered by the reader when interpreting our results. However, as the PFI data

used here is designed to account for meteorological change, and the only major observed change in PFI occurred in 2015 following the Gorkha earthquake, it is considered a fair assumption that our overall results and conclusions are not impacted by this data issue.

Over the past 30 yr, land use in Nepal has changed, with a review by Paudel et al. (2016) showing that the main changes are increases in the area of cropland and urban land, and small decreases in forest and permanent snow/ice/glaciers. However, accurately quantifying land use change, particularly for the periods before high resolution satellites, is challenging, with different studies showing differing magnitudes and types of land use change through time (B. Paudel et al., 2016). Furthermore, to our knowledge, there are no freely available land use products for our study region going back to the 1980s. As such, in this study, we used the ESA Glob-cover product, which as outlined in Section 3.5, gives the land use classifications as they were in 2009 at a 300×300 m spatial resolution. This raises the question of whether any land use misclassifications resulting from the static use of 2009 land use data has impacted our results.

The most temporally dynamic landuse product available is the MODIS dataset, which estimates global land cover between 2001 and 2018. As such, to assess if the methodology used here will have affected the results, the MODIS IGBP (International Geosphere-Biosphere Program) classifications for the landslide data from 2001 to 2018 are extracted. By grouping the MODIS classifications into the same broad categories as used for the ESA product, the number of landslides in each year from 2001 to 2018 that changed classifications relative to 2009 are counted. On average, less than 9% of landslides in the years 2001–2018 changed classification relative to 2009. Furthermore, the chi-square analysis for land use for the years 2001–2018 is repeated using the annual MODIS classification. If misclassification error in our approach is the cause of the variations observed in Figure 4b, then we would expect the repeated chi-square analysis with temporally variable land use data to now show a much higher level of stability. However, as shown in Figure S3 this is not the case, with significant variability between years, particularly in 2002, 2015 and 2017, as shown by the original analysis. This thus supports our assumption that the observed variability is dominantly caused by true physical processes, rather than data misclassification. Finally, as shown in Figures 5p and S2, most land use categories were rarely selected by the LASSO as being important. Furthermore, cropland, which was the most commonly selected landuse category, was still only selected consistently in 4 of the 12 modeled years. Figure 5p also shows that the cropland results are very consistent between years, in all but 1993, where cropland is suddenly defined as making landslides less likely. Unfortunately, the MODIS data does not cover this period, however, there are two possibilities for why the 1993 results for cropland changes. One, is that it is due to the use of stationary landuse data, that is, in 1993 there was some sudden change in cropland distributions not accounted for by the data used here that caused the observed change. Or two, that it is due to some other true physical process that shifted landslide distributions in that year to be less likely in croplands. Literature investigations reveal no evidence to support explanation one. Furthermore, if there was a sudden increase in cropland in 1993 that changed the results, then this change would have to reverse by 2000, when the results return to as they were before 1993. Again, literature investigations find no evidence for such a change happening. The other explanation for the 1993 cropland result is that it is due to some physical process that dramatically shifted landslide distributions in that year. It is known that 1993 experienced an extreme cloud outburst event (e.g., Dhital, 2003), which changed landslide distributions across most control factors investigate (Figures 4–5). This supports the assumption that the observed change in 1993 is attributable to a physical process, rather than data misclassification. However, it is important to acknowledge this limitation and to consider it alongside subsequent discussions.

6.2. Landslide Spatial Distributions

The first objective of this paper was to quantify whether landslide characteristics and spatial distributions varied through time. Overall, our K-S and chi-square results show that the landslide distributions for most control factors vary through time, with only profile and planform curvature, SSP and distance to channels having >70% similarity across the 29 mapped years. Most concerning, some of the control factors routinely used in landslide susceptibility models such as slope, elevation, geology, and landuse showed some of the greatest temporal variations. Of these, slope was the only factor to have similar distributions in >50% of years at all alpha values (0.01, 0.005, and 0.001), with geology having similar distributions in <30% (Figure 3).

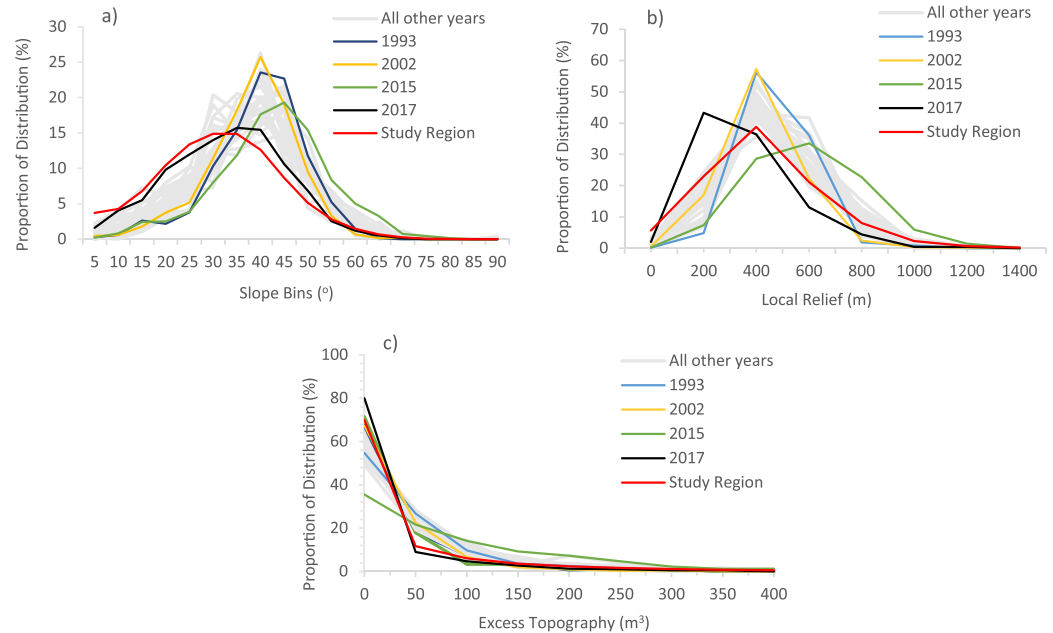


Figure 8. Proportion of the study region, all landslide data and landslide data per time slice that fall within given bins of (a) slope, (b) local relief, and (c) excess topography.

These results highlight that temporal variations in landslide spatial distributions exist even for landslide distributions across the same region and trigger events (in this case monsoonal rainfall). This result is similar to those of Kinney et al. (2021) who found significant changes in the characteristics and distributions of coseismic and monsoon-triggered landslides in Nepal across the period 2014–2018. But what causes these changes? It is apparent that in the results of this article, many of the most significant changes in landslide distributions occur in years coincident with extreme events, most notably with cloud outburst events 1993 and 2002, severe flooding in 2017 and the Gorkha earthquake in 2015. Despite some studies showing that extreme rainfall does not necessarily equate to increased landslides (e.g., Saito et al., 2014), it is reasonable to hypothesize that these extreme events can explain the observed changes in spatial distributions. To assess this hypothesis, we take a more detailed look at each of these extreme years using the results from Section 5.1. These new analyses show the proportions of landslides in each of the extreme years across several key control factors (slope, local relief and excess topography), as well as the proportions of all other years and the study region across those control factors (Figures 8a–8c).

In the case of 1993, the atypically narrow distributions for slope and local relief (Figures 8a and 8b) show that landslide locations shifted to steeper than normal slopes of 35–55° and higher than normal reliefs of 300–750 m. This likely reflects the localized nature of rainfall across hillslopes in the Mahabharat Mountains that are dominated by these topographic characteristics (Lavé & Avouac, 2001). In terms of excess topography, 1993 was one of the only years to have a negative (−0.1) regression coefficient (Figure 5k), suggesting that landslides were more likely in the absence of excess topography. This is geomorphologically sensible, as excess topography is likely to be expressed in the landscape as convex topography, yet it is known that intense rainfall is more likely to accumulate and cause high pore-pressures in concave topography (Chang et al., 2007; Kayastha, 2012). Landslides in 1993 also observed the largest selection percentage and regression coefficient for peak monthly rainfall (Figure 5n), further highlighting the importance of intense rainfall in controlling the distributions observed in this year.

In 2002 the distributions of landslides with slope were relatively average (Figure 8a). However, the local relief distribution shifted to higher than average relief's of 250–600 m (Figure 8b), which, like the 1993 event, is probably a consequence of the rainfall location, which occurred between the Siwalik Hills and the Mahabharat mountains where these reliefs dominate. Furthermore, landslides in 2002 closely followed the distribution of the study regions for excess topographies of up to 50 m, but also revealed that landslides

were slightly less likely than the average at excess topographies of 50–100 m (Figure 8c). Again, this is geomorphically feasible, as it is expected that landslides triggered by intense rainfall will preferentially occur at lower slopes with concave topographies. However, as indicated by the regression modeling, this was a less-dominant characteristic of landslides in 2002 as it was for landslides in 1993. Interestingly, 2002 was also characterized with a more dominant than usual regression coefficient result for distance to roads (Figure 5j), with a selection percentage of near 100% and a coefficient of -0.3 compared to near 0 for all other years. This is likely because the 2002 cloud outburst storms occurred across the Kathmandu valley and adjacent regions that are some of the most densely populated and urbanized in Nepal. As this region has significant urbanization, the impact of roads on landsliding is more apparent relative to other years where landsliding was distributed more homogeneously across regions with and without significant urbanization. The 2002 landslides were also found to be less likely in forested regions (Figure S2i), which is expected as deep-rooted trees can increase slope material cohesion and reduce pore water pressures via transpiration (Sidle & Bogaard, 2016).

The regression modeling shows that, with the exception of excess topography, 1993 and 2002 had similar topographic and hydrological distributions across factors including slope, relief, curvatures, peak rainfall, PFI, and elevation. However, these years had notably different distributions across geological and landuse units (Figures 4b, 4c, and S2a–S2n). This likely reflects the simple fact that the two outburst storms occurred over regions with slightly different landuse and geologies. Indeed, the region impacted by the 1993 storm was composed predominantly of forest and shrub/grass overlying granite/gneisses and schists, whilst the 2002 storm impacted a similar region to 1993 as well as regions dominated by cropland and dolomite. This highlights that similar events can have very different impacts on regression modeling coefficients for spatially heterogeneous control factors, even if those events are partially coincident as in this case. As such, when modeling future landslide occurrence, unless the region to be predicted is entirely coincident with the region in which the training data were obtained, it is important to keep in mind that regression coefficients for spatially heterogeneous control factors may be sub-optimal.

In the case of 2017, which was impacted by severe flooding, the landslide slope distribution closely follows the distribution the slope configuration of the landscape (Figure 8a), suggesting that, unusually, landslides in this case were not more likely to occur at any given slope range. In terms of local relief, 2017 was very different to all other years, showing that landslides were more likely where the relief is 0–400 m and less likely at reliefs of 400–1,000 m (Figure 8b). This may be because landslides triggered in this monsoon season were potentially more likely to have been influenced by undercutting of hillslopes due to turbulent flood waters and elevated flood water levels. This is in contrast with other years where triggering occurred in regions with the typical combination of higher slopes and higher rainfall. The excess topography distribution for 2017 corroborates the observation that landslides in this year occurred lower on slopes, as it shows that landslides were less likely than average to occur at higher excess topographies of 50–100 m (Figure 8c). In 2017 also saw the fourth highest regression coefficient for peak rainfall (Figure 5n), highlighting that intense rainfall remained a key determinant in the distributions of 2017, even if more of the landslides were triggered by floodwater. Furthermore, landslides in 2017 were much more likely in regions with cropland (Figure 5p), which is unsurprising given that poorly managed cultivation systems are known to increase infiltration rates and reduce soil cohesion, thus increasing pore water pressures whilst reducing regolith shear strength (Alexander, 1992).

In 2015, Figures 8a–8c, highlight that distributions of landslides with slope, local relief, and excess topography were all significantly different to the average. Specifically, slope distributions shift from an average of 35° – 60° to 45° – 70° , relief distributions shift from an average of 250–800 to 500–1200 m, and excess topography distributions shift from an average of 0–100 to 50–300 m^3 . Furthermore, the “distance to channels” factor had a much higher (90%) selection percentage compared to most years, with a negative regression coefficient of -0.075 . This suggests that in 2015, landslides were less likely to occur near channels, to the extent that this became an important factor for overall prediction of landslides in that year. These changes in topographic distributions are likely due to the topographic amplification of earthquake strong ground motion (e.g., Von Specht et al., 2019), which is often amplified near ridges and other topographic excesses (Meunier et al., 2008). The loading of hillslopes by ground motion has been observed via laboratory experiments to impact hillslope stability (Brain et al., 2017), with field observations following multiple earthquake

events showing that rainfall triggered landslides following earthquakes occur at higher rates and in different topographic locations (Hovius et al., 2011; C. W. Lin et al., 2006; Marc et al., 2019). As such, the changing topographic distributions of monsoon triggered landslides observed here during 2015 are likely a consequence of the topographic damage signature of the Gorkha earthquake. It is also observed that the regression coefficients for PFI and snow/ice cover in 2015 had much greater negative magnitudes than in other years. This is likely because the 2015 earthquake occurred in the Greater Himalaya, so the damage signature of the earthquake was in a region with pervasive snow/ice and permafrost. Consequently, the negative influence of snow/ice and permafrost on rainfall triggered landslide occurrence was more apparent in 2015 than it is in a typical monsoon season, where more landsliding occurs in the south where there is less pervasive snow/ice and permafrost. This highlights the importance of snow/ice cover and permafrost in preventing failures within landscapes. This is a salient point, as it is increasingly reported that snow/ice and permafrost cover within the Himalaya is reducing due to a warming climate (Gruber et al., 2017; Haeberli et al., 2017), thus representing a potential cause of increased future landslide activity. Overall, these results highlight that the transient topographic signature of earthquakes, as well as any large-scale time-dependent changes in permafrost distribution and snow/ice cover, should be considered in future landslide susceptibility modeling, highlighting the need to move toward susceptibility modeling methodologies that are more dynamic and time-dependent.

6.3. Impacts on Landslide Susceptibility Modeling

Our results clearly show that landslide distributions vary significantly through time, particularly in years impacted by extreme events. The second objective of this paper was to quantify the impacts of this variation on the accuracy and hindcasting ability of susceptibility models. As outlined in Section 4.2, this paper only considers one susceptibility approach, BLR modeling, so the wider applicability of our results for other susceptibility approaches requires further study.

Overall, we find that the years impacted by extreme events in 1993, 2002, 2015, and 2017 produced the worst performing susceptibility models, and were the hardest years for other models to hindcast. This suggests that the observed variations in individual landslide control factor distributions do have significant impacts on resulting susceptibility models. This suggests that BLR-based landslide susceptibility modeling needs to move away from time-independence, which assumes landslide distributions are static, toward more dynamic or time-dependent modeling that can account for expected or unexpected temporal variations in landslide distributions, particularly following extreme events. This suggestion follows a growing number of similar recommendations in the literature for landslide susceptibility to be considered temporally dynamic (e.g., Gorsevski et al., 2006; Lombardo et al., 2020; Meusburger & Alewell, 2009; Ozturk et al., 2021).

The growing evidence for the need for temporally dynamic susceptibility modeling raises the question of whether sufficient modeling techniques exist to account for temporal variance and, if so, why they are not more widely applied. As outlined by Lombardo et al. (2020), the main limitation of developing robust space-time landslide susceptibility models is the lack of available accurate multi-temporal landslide data, rather than the availability of complex statistical modeling tools. Indeed, Lombardo et al. (2020) present a regression based model that is, capable of accounting for latent temporal effects between consecutive inventories. Such methods should be capable of accounting for the more predictable observations here, such as the post-earthquake changes associated with topographic damage. This suggests that structured temporal variations could be accounted for provided that sufficient data exists to observe and identify any structure. Indeed, an example of structural temporal variation being accounted for within a susceptibility model can be seen in the path-dependent modeling of Samia et al. (2018). Using a ~60 yr multi-temporal landslide database for Italy, Samia et al. (2017) observed that landsliding was transiently time-dependent, as controlled by the time-limited heritage effects of previous landslide locations. They then developed time-dependent models that accounted for this structured transient change (Samia et al., 2018, 2020). However, whilst the work of Samia et al. (2018, 2020) and Lombardo et al. (2020) nicely demonstrates that structured temporal changes can be modeled given sufficiently long training data, as noted by Korup and Stolle (2014), it is less clear whether more stochastic temporal landslide processes, such as some of those observed here for individual control factors (e.g., Figures 4a–4f), can be modeled using existing statistical methods.

As such, a simpler solution to this problem could be the use of more than one inventory for a given prediction or hindcast. For example, Kritikos et al. (2015) found that for the spatial application of fuzzy logic methods, using two inventories from different locations allowed accurate modeling of landslides in a third location. This raises the question of whether, using landslide inventories for two separate time slices could allow accurate modeling of a third time slice. This leads to the final objective of this paper, which was to investigate the length of historical landslide inventory (or number of combined seasonal inventories) required to train a BLR susceptibility model that can be applied reliably through time.

As shown in Figure 6 single season inventories from extreme years do not offer consistent and reliable prediction (or hindcasting) through time. However, as shown in Figures 7a–7j, as landslide data from an extreme season is combined with increasingly more seasons of data, model accuracy increases rapidly as the number of seasons increases from 2 to 6 yr. This increase then begins to saturate as the numbers of seasons increases beyond 6–8 yr. This result is similar to that of Ozturk et al. (2021) who found that the accuracy of a logistic regression based susceptibility model saturated after 0.01% of the study region had failed and was used to train the model. The reason for this saturation likely relates to the averaging of landslide distributions. If a single inventory was impacted by some process that makes the landslides within it have atypical distributions (e.g., as seen here, earthquake preconditioning shifting landslides to higher-than-normal excess topographies, or storms shifting landslides to higher slopes but lower excess topographies) then a model developed from just that year will only be capable of predicting landslides with similarly atypical landslide occurrences. However, if more and more “typical” landslide data are used to develop BLR susceptibility model alongside that extreme year, then the resulting model will shift toward being applicable to the average landslide distributions for that region. However, as shown by Figure 7h, this only works when predicting a future year that is, itself typical. When predicting a future year that is, actually atypical (i.e., another extreme year), then it is possible that a model based on the longer term average landslide distributions will not be the best model to predict that year.

This highlights a fundamental problem in landslide susceptibility modeling, which is modeling blind of physical processes. This paper shows that models developed using single “normal season” (i.e., those impacted by a known process such as typical monsoonal rainfall) can consistently hindcast other similarly “normal” seasons, but that models developed from seasons impacted by another processes such as extreme rainfall cannot unless combined with six to eight seasons of other landslide data. Conversely, a model developed from a single season impacted by an extreme process might actually be well capable of predicting a similar extreme future season, in which case saturation with landslide data from normal seasons will make a model worse. Therefore, accurate susceptibility modeling (at least BLR modeling) requires that models developed from particular causal mechanism are only used to predict or hindcast that same mechanism. This finding is clearly pertinent to the Himalaya, but also has broader relevance to any region with multiple coincident landslide drivers. As such, future work should focus on investigating this problem in other tectonic and climatic regimes.

7. Conclusions

Overall, our results show that the spatial distributions of monsoon-triggered landslides varies significantly through time, particularly in response to cloud outburst events in 1993 and 2002, flooding in 2017 and earthquake preconditioning following the 2015 Gorkha earthquake. We find that the topographic damage signature of Gorkha earthquake preconditioning shifted 2015 monsoon-triggered landslides to higher slopes (45–70°), reliefs (500–1,200 m) and excess topographies (50–300 m³). Cloud outburst events in 1993 and 2002 also shifted landslides to higher-than-average slopes (35–55° for 1993) and reliefs (250–600 m in 2002; 300–750 m in 1993). However, in contrast to earthquake preconditioning, regression modeling suggests a slight tendency for landslides triggered by extreme rainfall to cluster at lower excess topographies. Flooding in 2017 shifted landslides to much lower-than-average reliefs (<400 m), slopes (0–35°), and excess topographies (50–100 m³). These variations are found to have significant impacts on BLR susceptibility modeling, with models trained on these extreme years unable to consistently predict the landslide occurrence in other years with sufficient accuracy. We suggest that there are two solutions to this: one, to use more complex time-dependent logistic regression based modeling such as that proposed by Lombardo et al. (2020), which is capable of exploiting latent temporal structure within a model provided the input

landslide data covers a sufficient time period to encompass any temporal variation. Two, to ensure that BLR susceptibility models are developed using historical or multi-seasonal landslide inventories with at least six to eight separate inventories of landslide data that average out any anomalous landslide distributions that may have occurred in periods impacted by extreme events, and thus provide more reliable prediction. Finally, regardless of the method used to deal with temporal variance, we conclude that it is vital to ensure that susceptibility modeling is not undertaken “process blind,” with susceptibility models only used to predict future landslide occurrences with similar causal mechanisms to the data used to train the model.

Data Availability Statement

The landslide inventory point data and geological data used throughout this manuscript are available as a.txt file and.tiff file from the National Geoscience Data Centre (NGDC) repository (item ID 14025): <https://webapps.bgs.ac.uk/services/ngdc/accessions/index.html?simpleText=nepal#item140245>; <https://webapps.bgs.ac.uk/services/ngdc/accessions/index.html?simpleText=nepal%20geology#item165666>.

Acknowledgments

This work was supported by the Natural Environment Research Council through the EnvEast Doctoral Training Partnership between the University of East Anglia and University of Plymouth [NE/L002582/1], with CASE-funding from AECOM. The authors would like to thank the editor, Noah Finnegan, associate editor, Matt Brain, reviewer, Ugur Ozturk, and two anonymous reviewers for their comments and suggestions that improved the quality of this manuscript.

References

- Arditiani, A., Kubota, T., & Shinohara, Y. (2018). Comparison of GIS-based landslide susceptibility models using frequency ratio, logistic regression, and artificial neural network in a tertiary region of Ambon, Indonesia. *Geomorphology*, 318, 101–111. <https://doi.org/10.1016/J.GEOMORPH.2018.06.006>
- Aleotti, P., & Chowdhury, R. (1999). Landslide hazard assessment: Summary review and new perspectives. *Bulletin of Engineering Geology and the Environment*, 58, 21–44. <https://doi.org/10.1007/s100640050066>
- Alexander, D. (1992). On the causes of landslides: Human activities, perception, and natural processes. *Environmental Geology and Water Sciences*, 20, 165–179. <https://doi.org/10.1007/BF01706160>
- Alexandridis, T. K., Cherif, I., Kalogeropoulos, C., Monachou, S., Eskridge, K., & Silleos, N. (2013). Rapid error assessment for quantitative estimations from Landsat 7 gap-filled images. *Remote Sensing Letters*, 4, 920–928. <https://doi.org/10.1080/2150704X.2013.815380>
- Alvioli, M., Marchesini, I., Reichenbach, P., Rossi, M., Ardizzone, F., Fiorucci, F., & Guzzetti, F. (2016). Automatic delineation of geomorphological slope units with r.slopeunits v1.0 and their optimization for landslide susceptibility modeling. *Geoscientific Model Development*, 9, 3975–3991. <https://doi.org/10.5194/gmd-9-3975-2016>
- Amato, G., Eisank, C., Castro-Camilo, D., & Lombardo, L. (2019). Accounting for covariate distributions in slope-unit-based landslide susceptibility models. A case study in the alpine environment. *Engineering Geology*, 260, 105237. <https://doi.org/10.1016/j.enggeo.2019.105237>
- Ao, M., Zhang, L., Dong, Y., Su, L., Shi, X., Balz, T., & Liao, M. (2020). Characterizing the evolution life cycle of the Sunkoshi landslide in Nepal with multi-source SAR data. *Scientific Reports*, 10, 17988. <https://doi.org/10.1038/s41598-020-75002-y>
- Ashouri, H., Hsu, K.-L., Sorooshian, S., Braithwaite, D. K., Knapp, K. R., Cecil, L. D., et al. (2015). PERSIANN-CDR: Daily precipitation climate data record from multisatellite observations for hydrological and climate studies. *Bulletin of the American Meteorological Society*, 96, 69–83. <https://doi.org/10.1175/BAMS-D-13-00068.1>
- Baeza, C., & Corominas, J. (2001). Assessment of shallow landslide susceptibility by means of multivariate statistical techniques. *Earth Surface Processes and Landforms*, 26, 1251–1263. <https://doi.org/10.1002/esp.263>
- Beck, H. E., Vergopolan, N., Pan, M., Levizzani, V., Van Dijk, A. I. J. M., Weedon, G. P., et al. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, 21, 6201–6217. <https://doi.org/10.5194/hess-21-6201-2017>
- Behling, R., Roessner, S., Golovko, D., & Kleinschmit, B. (2016). Derivation of long-term spatiotemporal landslide activity—A multi-sensor time series approach. *Remote Sensing of Environment*, 186, 88–104. <https://doi.org/10.1016/j.rse.2016.07.017>
- Bennett, G. L., Miller, S. R., Roering, J. J., & Schmidt, D. A. (2016). Landslides, threshold slopes, and the survival of relict terrain in the wake of the Mendocino Triple Junction. *Geology*, 44, 363–366. <https://doi.org/10.1130/G37530.1>
- Blöthe, J. H., Korup, O., & Schwanghart, W. (2015). Large landslides lie low: Excess topography in the Himalaya-Karakoram ranges. *Geology*, 43, 523–526. <https://doi.org/10.1130/G36527.1>
- Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., & Ramos Perez, J. (2011). *GLOBCOVER 2009 products description and validation report*. Université Catholique de Louvain (UCL) & European Space Agency.
- Brain, M. J., Rosser, N. J., & Tunstall, N. (2017). The control of earthquake sequences on hillslope stability. *Geophysical Research Letters*, 44, 865–872. <https://doi.org/10.1002/2016GL071879>
- Braun, A., Fernandez-Steege, T., Havenith, H. B., & Torgoev, A. (2015). Landslide susceptibility mapping with data mining methods—A case study from Maily-Say, Kyrgyzstan. In *Engineering geology for society and territory: Landslide processes* (Vol. 2, pp. 995–998). Springer International Publishing. https://doi.org/10.1007/978-3-319-09057-3_174
- Cascini, L. (2008). Applicability of landslide susceptibility and hazard zoning at different scales. *Engineering Geology*, 102, 164–177. <https://doi.org/10.1016/j.enggeo.2008.03.016>
- Chang, K.-T., Chiang, S.-H., & Hsu, M.-L. (2007). Modeling typhoon- and earthquake-induced landslides in a mountainous watershed using logistic regression. *Geomorphology*, 89, 335–347. <https://doi.org/10.1016/J.GEOMORPH.2006.12.011>
- Chauhan, R., & Thakuri, S. (2017). Periglacial environment in Nepal Himalaya: Present contexts and future prospects. *Nepal Journal of Environmental Science*, 5, 35–40. <https://doi.org/10.3126/njes.v5i0.22713>
- Conforti, M., Robustelli, G., Muto, F., Critelli, S., Conforti, M., Robustelli, G., et al. (2012). Application and validation of bivariate GIS-based landslide susceptibility assessment for the Vitraro river catchment (Calabria, south Italy). *Natural Hazards*, 61, 127–141. <https://doi.org/10.1007/s11069-011-9781-0>

- Craddock, W. H., Burbank, D. W., Bookhagen, B., Gabet, E. J., Burbank, D. W., Bookhagen, B., & Gabet, E. J. (2007). Bedrock channel geometry along an orographic rainfall gradient in the upper Marsyandi River valley in central Nepal. *Journal of Geophysical Research*, 112, 3007. <https://doi.org/10.1029/2006JF000589>
- Dahal, R. K., & Hasegawa, S. (2008). Representative rainfall thresholds for landslides in the Nepal Himalaya. *Geomorphology*, 100, 429–443. <https://doi.org/10.1016/J.GEOMORPH.2008.01.014>
- Defourny, P., Schouten, L., Bartalev, S., Bontemps, S., Caccetta, P., De Wit, A., et al. (2009). *Accuracy assessment of a 300 m global land cover map: The GlobCover experience*. CIRAD-Guyane-Université Laval.
- Densmore, A. L., & Hovius, N. (2000). Topographic fingerprints of bedrock landslides. *Geology*, 28, 371–374. [https://doi.org/10.1130/0091-7613\(2000\)028<0371:tfobl>2.3.co;2](https://doi.org/10.1130/0091-7613(2000)028<0371:tfobl>2.3.co;2)
- Dhital, M. R. (2003). Causes and consequences of the 1993 debris flows and landslides in the Kulekhani watershed, central Nepal. In D. Rickenmann & C. L. Chen (Eds.), *Proceedings of 3rd International Conference on Debris-Flow Hazards Mitigation: Mechanics, prediction and assessment* (pp. 931–942). Millpress.
- Dilley, M. (2005). *Natural disaster hotspots: A global risk analysis*. World Bank.
- Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., & Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land-use planning on behalf of the JTC-1 Joint Technical Committee on Landslides and Engineered Slopes. *Engineering Geology*, 102, 99–111. <https://doi.org/10.1016/j.enggeo.2008.03.014>
- Fiorucci, F., Cardinali, M., Carlà, R., Rossi, M., Mondini, A. C., Santurri, L., et al. (2011). Seasonal landslide mapping and estimation of landslide mobilization rates using aerial and satellite images. *Geomorphology*, 129, 59–70. <https://doi.org/10.1016/j.geomorph.2011.01.013>
- French, H., Heginbottom, J., Johnston, G., Ladanyi, B., Sego, D., & van Everdingen, R. (1988). *Glossary of permafrost and related ground-ice terms*. National Research Council of Canada.
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33, 1–22. <https://doi.org/10.18637/jss.v033.i01>
- Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18, 2161–2181. <https://doi.org/10.5194/nhess-18-2161-2018>
- Fukui, K., Fujii, Y., Ageta, Y., & Asahi, K. (2007). Changes in the lower limit of mountain permafrost between 1973 and 2004 in the Khumbu Himal, the Nepal Himalaya. *Global and Planetary Change*, 55, 251–256. <https://doi.org/10.1016/j.gloplacha.2006.06.002>
- Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., & Reichenbach, P. (2008). Comparing landslide inventory maps. *Geomorphology*, 94, 268–289. <https://doi.org/10.1016/j.geomorph.2006.09.023>
- Gautam, D., & Dong, Y. (2018). Multi-hazard vulnerability of structures and lifelines due to the 2015 Gorkha earthquake and 2017 central Nepal flash flood. *Journal of Building Engineering*, 17, 196–201. <https://doi.org/10.1016/j.jobe.2018.02.016>
- Goetz, J. N., Guthrie, R. H., & Brenning, A. (2011). Integrating physical and empirical landslide susceptibility models using generalized additive models. *Geomorphology*, 129, 376–386. <https://doi.org/10.1016/j.geomorph.2011.03.001>
- Gorsevski, P. V., Gessler, P. E., Boll, J., Elliot, W. J., & Foltz, R. B. (2006). Spatially and temporally distributed modeling of landslide susceptibility. *Geomorphology*, 80, 178–198. <https://doi.org/10.1016/j.geomorph.2006.02.011>
- Government of Nepal. (2017). *Nepal flood 2017: Post flood recovery needs assessment*. National Planning Commission.
- Gruber, S. (2012). Derivation and analysis of a high-resolution estimate of global permafrost zonation. *Cryosphere*, 6, 221–233. <https://doi.org/10.5194/tc-6-221-2012>
- Gruber, S., Fleiner, R., Guegan, E., Panday, P., Schmid, M. O., Stumm, D., et al. (2017). Review article: Inferring permafrost and permafrost thaw in the mountains of the Hindu Kush Himalaya region. *Cryosphere*, 11, 81–99. <https://doi.org/10.5194/tc-11-81-2017>
- Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K.-T. (2012). Landslide inventory maps: New tools for an old problem. *Earth-Science Reviews*, 112, 42–66. <https://doi.org/10.1016/J.EARSCIREV.2012.02.001>
- Guzzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., & Galli, M. (2006). Estimating the quality of landslide susceptibility models. *Geomorphology*, 81, 166–184. <https://doi.org/10.1016/J.GEOMORPH.2006.04.007>
- Haeberli, W., Schaub, Y., & Huggel, C. (2017). Increasing risks related to landslides from degrading permafrost into new lakes in de-glaciating mountain ranges. *Geomorphology*, 293, 405–417. <https://doi.org/10.1016/j.geomorph.2016.02.009>
- Hastie, T., & Qian, J. (2014). *Glmmnet vignette* [WWW Document]. Retrieved from https://web.stanford.edu/~hastie/glmmnet/glmmnet_alpha.html
- Hovius, N., Meunier, P., Lin, C.-W., Chen, H., Chen, Y.-G., Dadson, S., et al. (2011). Prolonged seismically induced erosion and the mass balance of a large earthquake. *Earth and Planetary Science Letters*, 304, 347–355. <https://doi.org/10.1016/J.EPSL.2011.02.005>
- Humanitarian Data Exchange. (2020). *Humanitarian OpenStreetMap Team (HOT)* [WWW Document]. Retrieved from <https://data.humdata.org/organization/hot>
- Jaiswal, P., van Westen, C. J., & Jetten, V. (2011). Quantitative assessment of landslide hazard along transportation lines using historical records. *Landslides*, 8, 279–291. <https://doi.org/10.1007/s10346-011-0252-1>
- Kayastha, P. (2012). Application of fuzzy logic approach for landslide susceptibility mapping in Garuwa sub-basin, East Nepal. *Frontiers of Earth Science*, 6, 420–432. <https://doi.org/10.1007/s11707-012-0337-8>
- Kinney, M. E., Rosser, N. J., Robinson, T. R., Densmore, A. L., Shrestha, R., Pujara, D. S., et al. (2021). Evolution of coseismic and post-seismic landsliding after the 2015 M_w 7.8 Gorkha Earthquake, Nepal. *Journal of Geophysical Research: Earth Surface*, 126, e2020JF005803. <https://doi.org/10.1029/2020JF005803>
- Kirchner, J. W., Finkel, R. C., Riebe, C. S., Granger, D. E., Clayton, J. L., King, J. G., & Megahan, W. F. (2001). Mountain erosion over 10 yr, 10 k.y., and 10 m.y. time scales. *Geology*, 29, 591–594. [https://doi.org/10.1130/0091-7613\(2001\)029<0591:MEQYK>2.0.CO;2](https://doi.org/10.1130/0091-7613(2001)029<0591:MEQYK>2.0.CO;2)
- Korup, O., & Stolle, A. (2014). Landslide prediction from machine learning. *Geology Today*, 30, 26–33. <https://doi.org/10.1111/gto.12034>
- Kritikos, T., Robinson, T. R., & Davies, T. R. H. (2015). Regional coseismic landslide hazard assessment without historical landslide inventories: A new approach. *Journal of Geophysical Research: Earth Surface*, 120, 711–729. <https://doi.org/10.1002/2014JF003224>
- Larsen, I. J., & Montgomery, D. R. (2012). Landslide erosion coupled to tectonics and river incision. *Nature Geoscience*, 5, 468–473. <https://doi.org/10.1038/ngeo1479>
- Lavé, J., & Avouac, J. P. (2001). Fluvial incision and tectonic uplift across the Himalaya of central Nepal. *Journal of Geophysical Research: Solid Earth*, 106, 26561–26591. <https://doi.org/10.1029/2001JB000359>
- Lee, C.-T., Huang, C.-C., Lee, J.-F., Pan, K.-L., Lin, M.-L., & Dong, J.-J. (2008). Statistical approach to storm event—Induced landslides susceptibility. *Natural Hazards and Earth System Sciences*, 8, 941–960. <https://doi.org/10.5194/nhess-8-941-2008>
- Lin, C. W., Liu, S. H., Lee, S. Y., & Liu, C. C. (2006). Impacts of the Chi-Chi earthquake on subsequent rainfall-induced landslides in central Taiwan. *Engineering Geology*, 86, 87–101. <https://doi.org/10.1016/j.enggeo.2006.02.010>

- Lin, L., Lin, Q., & Wang, Y. (2017). Landslide susceptibility mapping on a global scale using the method of logistic regression. *Natural Hazards and Earth System Sciences*, 17, 1411–1424. <https://doi.org/10.5194/nhess-17-1411-2017>
- Lombardo, L., & Mai, P. M. (2018). Presenting logistic regression-based landslide susceptibility results. *Engineering Geology*, 244, 14–24. <https://doi.org/10.1016/j.enggeo.2018.07.019>
- Lombardo, L., Opitz, T., Ardizzone, F., Guzzetti, F., & Huser, R. (2020). Space-time landslide predictive modelling. *Earth-Science Reviews*, 209, 103318. <https://doi.org/10.1016/j.earscirev.2020.103318>
- Malamud, B. D., Turcotte, D. L., Guzzetti, F., & Reichenbach, P. (2004). Landslide inventories and their statistical properties. *Earth Surface Processes and Landforms*, 29, 687–711. <https://doi.org/10.1002/esp.1064>
- Marc, O., Behling, R., Andermann, C., Turowski, J. M., Illien, L., Roessner, S., & Hovius, N. (2019). Long-term erosion of the Nepal Himalaya by bedrock landsliding: The role of monsoons, earthquakes and giant landslides. *Earth Surface Dynamics*, 7, 107–128. <https://doi.org/10.5194/esurf-7-107-2019>
- Marc, O., & Hovius, N. (2015). Amalgamation in landslide maps: Effects and automatic detection. *Natural Hazards and Earth System Sciences*, 15, 723–733. <https://doi.org/10.5194/nhess-15-723-2015>
- Marc, O., Hovius, N., Meunier, P., Uchida, T., & Hayashi, S. (2015). Transient changes of landslide rates after earthquakes. *Geology*, 43, 883–886. <https://doi.org/10.1130/G36961.1>
- Marjanović, M. (2013). Comparing the performance of different landslide susceptibility models in ROC space. In *Landslide science and practice: Landslide inventory and susceptibility and hazard zoning* (pp. 579–584). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-642-31325-7_76
- Martha, T. R., Kerle, N., van Westen, C. J., Jetten, V., & Vinod Kumar, K. (2012). Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 105–119. <https://doi.org/10.1016/j.isprsjprs.2011.11.004>
- Mcadoo, B. G., Quak, M., Gnyawali, K. R., Adhikari, B. R., Devkota, S., Rajbhandari, P. L., & Sudmeier-Rieux, K. (2018). Roads and landslides in Nepal: How development affects environmental risk. *Natural Hazards and Earth System Sciences*, 18, 3203–3210. <https://doi.org/10.5194/nhess-18-3203-2018>
- Meunier, P., Hovius, N., & Haines, J. A. (2008). Topographic site effects and the location of earthquake induced landslides. *Earth and Planetary Science Letters*, 275, 221–232. <https://doi.org/10.1016/j.epsl.2008.07.020>
- Meusburger, K., & Alewell, C. (2009). On the influence of temporal change on the validity of landslide susceptibility maps. *Natural Hazards and Earth System Sciences*, 9(4), 1495–1507. <https://doi.org/10.5194/nhess-9-1495-2009>
- Molnar, P., & England, P. (1990). Late Cenozoic uplift of mountain ranges and global climate change: Chicken or egg? *Nature*, 346, 29–34. <https://doi.org/10.1038/346029a0>
- Murton, G. (2017). Bordering spaces, practising borders: Fences, roads and reorientations across a Nepal–China borderland. *South Asia: Journal of South Asian Studies*, 40, 239–255. <https://doi.org/10.1080/00856401.2017.1292616>
- Ozturk, U., Pittore, M., Behling, R., Roessner, S., Andreani, L., & Korup, O. (2021). How robust are landslide susceptibility estimates? *Landslides*, 18, 681–695. <https://doi.org/10.1007/s10346-020-01485-5>
- Palau, R. M., Hürlimann, M., Berenguer, M., & Sempere-Torres, D. (2020). Influence of the mapping unit for regional landslide early warning systems: Comparison between pixels and polygons in Catalonia (NE Spain). *Landslides*, 17, 2067–2083. <https://doi.org/10.1007/s10346-020-01425-3>
- Park, H. J., Jang, J. Y., & Lee, J. H. (2019). Assessment of rainfall-induced landslide susceptibility at the regional scale using a physically based model and fuzzy-based Monte Carlo simulation. *Landslides*, 16, 695–713. <https://doi.org/10.1007/s10346-018-01125-z>
- Parker, R. N., Hancox, G. T., Petley, D. N., Massey, C. I., Densmore, A. L., & Rosser, N. J. (2015). Spatial distributions of earthquake-induced landslides and hillslope preconditioning in the northwest South Island, New Zealand. *Earth Surface Dynamics*, 3, 501–525. <https://doi.org/10.5194/esurf-3-501-2015>
- Paudel, B., Zhang, Y.-L., Li, S.-C., Liu, S. L., Wu, X., Liu, N. R., & Khanal, N. R. (2016). Review of studies on land use and land cover change in Nepal. *Journal of Mountain Science*, 13, 643–660. <https://doi.org/10.1007/s11629-015-3604-9>
- Paudel, P. P., Omura, H., Kubota, T., & Morita, K. (2003). Landslide damage and disaster management system in Nepal. *Disaster Prevention and Management*, 12, 413–419. <https://doi.org/10.1108/09653560310507235>
- Petley, D. N., Rosser, A. N. J., Dunning, A. S. A., Oven, A. K., Mitchell, A. W. A., Hearn, G. J., et al. (2007). Trends in landslide occurrence in Nepal. *Natural Hazards*, 43, 23–44. <https://doi.org/10.1007/s11069-006-9100-3>
- Rahaman, W., Singh, S. K., Sinha, R., & Tandon, S. K. (2009). Climate control on erosion distribution over the Himalaya during the past ~100 ka. *Geology*, 37, 559–562. <https://doi.org/10.1130/G25425A.1>
- Rana, K., Ozturk, U., & Malik, N. (2021). Landslide geometry reveals its trigger. *Geophysical Research Letters*, 48, e2020GL090848. <https://doi.org/10.1029/2020GL090848>
- Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., & Guzzetti, F. (2018). A review of statistically-based landslide susceptibility models. *Earth-Science Reviews*, 180, 60–91. <https://doi.org/10.1016/j.earscirev.2018.03.001>
- Roback, K., Clark, M. K., West, A. J., Zekkos, D., Li, G., Gallen, S. F., et al. (2018). The size, distribution, and mobility of landslides caused by the 2015 M_w 7.8 Gorkha earthquake, Nepal. *Geomorphology*, 301, 121–138. <https://doi.org/10.1016/j.geomorph.2017.01.030>
- Robinson, T. R., Rosser, N. J., Densmore, A. L., Williams, J. G., Kinney, M. E., Benjamin, J., & Bell, H. J. A. (2017). Rapid post-earthquake modelling of coseismic landslide intensity and distribution for emergency response decision support. *Natural Hazards and Earth System Sciences*, 17, 1521–1540. <https://doi.org/10.5194/nhess-17-1521-2017>
- Romer, C., & Ferentinou, M. (2016). Shallow landslide susceptibility assessment in a semiarid environment—A quaternary catchment of KwaZulu-Natal, South Africa. *Engineering Geology*, 201, 29–44. <https://doi.org/10.1016/j.enggeo.2015.12.013>
- Rossi, M., & Malamud, B. D. (2014). *LAMPRE, Prototype SW for determination of landslide statistics from inventory maps*.
- Sabatidakis, N., Koukis, G., Vassiliades, E., & Lainas, S. (2013). Landslide susceptibility zonation in Greece. *Natural Hazards*, 65, 523–543. <https://doi.org/10.1007/s11069-012-0381-4>
- Saito, H., Korup, O., Uchida, T., Hayashi, S., & Oguchi, T. (2014). Rainfall conditions, typhoon frequency, and contemporary landslide erosion in Japan. *Geology*, 42, 999–1002. <https://doi.org/10.1130/G35680.1>
- Samia, J., Temme, A., Bregt, A., Wallinga, J., Guzzetti, F., & Ardizzone, F. (2020). Dynamic path-dependent landslide susceptibility modelling. *Natural Hazards and Earth System Sciences*, 20, 271–285. <https://doi.org/10.5194/nhess-20-271-2020>
- Samia, J., Temme, A., Bregt, A., Wallinga, J., Ardizzone, F., & Rossi, M. (2017). Do landslides follow landslides? Insights in path dependency from a multi-temporal landslide inventory. *Landslides*, 14, 547–558. <https://doi.org/10.1007/s10346-016-0739-x>
- Samia, J., Temme, A., Bregt, A. K., Wallinga, J., Stuiver, J., Guzzetti, F., et al. (2018). Implementing landslide path dependency in landslide susceptibility modelling. *Landslides*, 15, 2129–2144. <https://doi.org/10.1007/s10346-018-1024-y>

- Sassa, K., & Canuti, P. (Eds.). (2009). *Landslides—Disaster risk reduction*. Springer. <https://doi.org/10.1007/978-3-540-69970-5>
- Schwanghart, W., & Scherler, D. (2014). Short Communication: TopoToolbox 2—MATLAB-based software for topographic analysis and modeling in Earth surface sciences. *Earth Surface Dynamics*, 2, 1–7. <https://doi.org/10.5194/esurf-2-1-2014>
- Sidle, R. C., & Bogaard, T. A. (2016). Dynamic earth system and ecological controls of rainfall-initiated landslides. *Earth-Science Reviews*, 159, 275–291. <https://doi.org/10.1016/j.earscirev.2016.05.013>
- Stanley, T., & Kirschbaum, D. B. (2017). A heuristic approach to global landslide susceptibility mapping. *Natural Hazards*, 87, 145–164. <https://doi.org/10.1007/s11069-017-2757-y>
- Stanley, T., Kirschbaum, D. B., Pascale, S., & Kapnick, S. (2020). Extreme precipitation in the Himalayan landslide hotspot. In *Advances in global change research* (pp. 1087–1111). Springer. https://doi.org/10.1007/978-3-030-35798-6_31
- Sulla-Menashe, D., & Friedl, M. A. (2018). *User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product*. Retrieved from https://lpdaac.usgs.gov/documents/101/MCD12_User_Guide_V6.pdf
- Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56, 79–107. <https://doi.org/10.1002/2017RG000574>
- Syvitski, J. P. M., & Kettner, A. (2011). Sediment flux and the Anthropocene. *Philosophical Transactions of the Royal Society A: Mathematical Physical and Engineering Sciences*, 369, 957–975. <https://doi.org/10.1098/rsta.2010.0329>
- Takai, N., Shigefuji, M., Rajaure, S., Bijukchhen, S., Ichianagi, M., Dhital, M. R., & Sasatani, T. (2016). Strong ground motion in the Kathmandu Valley during the 2015 Gorkha, Nepal, earthquake. *Earth, Planets and Space*, 68, 10. <https://doi.org/10.1186/s40623-016-0383-7>
- Thapa, S., Shrestha, A., Lamichhane, S., Adhikari, R., & Gautam, D. (2020). Catchment-scale flood hazard mapping and flood vulnerability analysis of residential buildings: The case of Khando River in eastern Nepal. *Journal of Hydrology: Regional Studies*, 30, 100704. <https://doi.org/10.1016/j.ejrh.2020.100704>
- Thi Ngo, P. T., Panahi, M., Khosravi, K., Ghorbanzadeh, O., Karimnejad, N., Cerda, A., & Lee, S. (2020). Evaluation of deep learning algorithms for national scale landslide susceptibility mapping of Iran. *Geoscience Frontiers*. <https://doi.org/10.1016/j.gsf.2020.06.013>
- Vakhshoori, V., & Zare, M. (2018). Is the ROC curve a reliable tool to compare the validity of landslide susceptibility maps? *Geomatics, Natural Hazards and Risk*, 9, 249–266. <https://doi.org/10.1080/19475705.2018.1424043>
- van Westen, C. J., Castellanos, E., & Kuriakose, S. L. (2008). Spatial data for landslide susceptibility, hazard, and vulnerability assessment: An overview. *Engineering Geology*, 102, 112–131. <https://doi.org/10.1016/j.enggeo.2008.03.010>
- Von Specht, S., Ozturk, U., Veh, G., Cotton, F., & Korup, O. (2019). Effects of finite source rupture on landslide triggering: The 2016 M_w 7.1 Kumamoto earthquake. *Solid Earth*, 10, 463–486. <https://doi.org/10.5194/se-10-463-2019>
- Wang, H. J., Xiao, T., Li, X. Y., Zhang, L. L., & Zhang, L. M. (2019). A novel physically-based model for updating landslide susceptibility. *Engineering Geology*, 251, 71–80. <https://doi.org/10.1016/j.enggeo.2019.02.004>
- Wang, X., Tolksdorf, V., Otto, M., & Scherer, D. (2021). WRF-based dynamical downscaling of ERA5 reanalysis data for High Mountain Asia: Toward a new version of the High Asia Refined analysis. *International Journal of Climatology*, 41, 743–762. <https://doi.org/10.1002/joc.6686>
- Yatagai, A., Kamiguchi, K., Arakawa, O., Hamada, A., Yasutomi, N., & Kitoh, A. (2012). Aphrodite constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of the American Meteorological Society*, 93, 1401–1415. <https://doi.org/10.1175/BAMS-D-11-00122.1>
- Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1, 3–14. <https://doi.org/10.1111/j.2041-210x.2009.00001.x>